Journal of Experimental Psychology

M

ARTHUR W. MELTON, Editor

HUMAN RESOURCES RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Robert M. Gagné, *Human Resources Research Center*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *Human Resources Research Center*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Edwin B. Newman, *Harvard University*
Kenneth W. Spence, *State University of Iowa*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

CONTENTS

Generalization of the Conditioned Galvanic Skin Response to Visual Stimuli: D. A. GRANT AND J. J. SCHILLER 309

Scotopic Area-Intensity Relations at Various Retinal Locations: A. J. RIOPELLE AND K. L. CHOW 314

Characteristics of the Muscle Tension Response to Paired Tones: D. W. VAN LIERE 319

Perceptual Judgment as a Function of Mental Set, Anchoring Point and Method of Judgment: C. W. HILL 325

The Effects of Various Kinds of Relevant Verbal Pretraining on Subsequent Motor Performance: D. E. MCALLISTER 329

Bilateral Reminiscence in Pursuit-Rotor Learning as a Function of Amount of First-Hand Practice and Length of Rest: M. R. ROCKWAY 337

Twenty Questions: An Information Analysis: A. W. BENDIG 345

An Effect of Repeated Conditioning-Extinction upon Operant Strength: D. H. BULLOCK AND W. C. SMITH 349

The Effect of Physical Constants of a Control on Tracking Performance: D. HOWLAND AND M. E. NOBLE 353

A Quantitative Approach to Figural "Goodness": J. HOCHBERG AND E. MCALISTER 361

Simultaneous Vision and Audition: The Comprehension of Prose Passages with Varying Levels of Difficulty: G. H. MOWBRAY 365

An Informational Analysis of Absolute Judgments of Loudness: W. R. GARNER 373

American Psychological Association

Vol. 46 No. 5 November 1953

DEC 1 - 1953

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc. The subscription rate per volume is $7.50, or $15.00 annually. Single copies are $1.50. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors who request them are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author, and one-half of the cost of tables, cuts, and formulas. Priority in publication is given to articles whose authors assume the full cost of publication.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Human Resources Research Center, Lackland Air Force Base, San Antonio, Texas.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 (4, Suppl.), 389–449. In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879.

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Journal of
Experimental Psychology

VOL. 46, No. 5 NOVEMBER, 1953

GENERALIZATION OF THE CONDITIONED GALVANIC SKIN RESPONSE TO VISUAL STIMULI [1]

DAVID A. GRANT AND JEROME J. SCHILLER
University of Wisconsin

The purpose of this experiment was to provide empirical data on the nature of a pure stimulus generalization gradient of the conditioned GSR. By pure generalization is meant generalization during which the process of discrimination is minimized. Operationally this reduces to investigating the magnitude of the response during the first test trial after conditioning. Subsequent test trials without reinforcement introduce discrimination training and hence interfere with pure generalization. It is, of course, recognized that generalization must presumably take place along a continuum of some sort and that in setting up this continuum in the past training of *S* there are all sorts of opportunities for secondary phenomena to develop (**17, 20**).

Stimulus generalization in one form or another occupies a central place in most learning theory (**10, 11**). In spite of this the empirical evidence on generalization in the relatively simple conditioning situation has been very confused (**23**), and the generalization following conditioning of the GSR has led to especially contradictory findings (**12, 13, 14, 15, 16, 18, 21, 23**). The desirability of clarifying the situation has been recognized generally, and other investigators have reported preliminary evidence on pure generalization of the conditioned GSR (**24**).

[1] This research was supported in part by the Graduate Research Committee of the Graduate School of the University of Wisconsin from funds appropriated by the State Legislature. The statistical transformations used were calculated by the Computing Service of the University of Wisconsin.

METHOD

Experimental conditions.—The experiment was carried out in a small sound-treated room. Within the room was placed most of the stimulating and recording equipment except for some relays which were housed in a soundproof compartment outside the room. Also within the room was a small masonite compartment in which *S* sat. The interior of the compartment was covered by white mat cardboard except for a glass screen about 3 ft. in front of *S*. The milk-glass screen was about 22 in. square. Two 60-w. bulbs in the rear in the booth provided an illumination of about 8 millilamberts on the walls and screen.

The *S*'s right palm rested on two GSR electrodes on the armrest of his classroom chair. The hand was held in place by a wrist clamp and a woven cloth strap which passed directly over the top of the wrist. On his left forearm were placed polished brass electrodes to deliver the electric shock which served as the UCS. A sound screen of about 40-db intensity level was provided by an electric fan placed on top of the booth. Aside from providing ventilation in the outer room the electric fan effectively masked minor operating sounds of the apparatus.

Stimuli.—The CS during training was a rectangle of light 12 in. high and 1 in. wide which

was projected on the center of the milk-glass screen. The generalization test stimuli were 1 in. wide and 9, 10, 11, 12, 13, 14, and 15 in. high, respectively. It was thought that by using relatively narrow rectangles area and intensity effects would be minimized (1). The duration of CS and generalization test stimuli was held constant at 1 sec.

The UCS was an electric shock delivered by a Muenzinger-Walz circuit (22). The intensity of the shock was adjusted so as to be unpleasant but not painful for *S*. The time relations between stimuli were controlled automatically by an electronic timer so that the UCS was delivered .5 sec. after the onset of the CS and lasted for .2 sec.

Recording system.—Nonpolarizable GSR electrodes (2, 21) were mounted .75 in. apart on a plastic block screwed to the arm of the chair. Thus both the electrodes were in contact with *S*'s palm so as to provide a relatively homogeneous GSR. A modified Greenwald (4) DC bridge passed an essentially constant current of about .09 ma. through *S*. The slight changes in voltage caused by resistance changes in the system were amplified and used to drive an ink-writing galvanometer that was calibrated to read in resistance change units. The initial level of *S*'s resistance before each trial could be read from the dials of the Greenwald circuit, and the change in resistance could be determined from the ink records by means of calibration curves.

Procedure.—After *S* was adjusted to the apparatus he was instructed to relax while the instruments warmed up. Five minutes were allowed for stabilization of *S*'s skin resistance, after which five adaptation trials were given with the CS. Three test shocks were then administered in order to set the level of the UCS for the individual *S*. The *S* was then told that a line would be projected on the screen and would be followed by a shock delivered to his left arm. He was told to lean back and relax, and that the nature of the experiment would be discussed at the close of the session. Thereafter *S* was given 25 training trials with the CS and UCS. Up to and including this period, the procedure was identical with all *S*s. At the end of the training trials 15 extinction trials were given with the appropriate generalization test stimulus. The generalization test stimulus was the same on all extinction trials for a given *S*. The interval between trials was varied unsystematically from 25 to 45 sec. for an average of about 35 sec.

Subjects.—One hundred and five men and women students who volunteered from the elementary psychology classes during the spring and fall terms of 1951 served as *S*s. The 105 *S*s were subdivided into seven groups of 15 each.[2] All groups were trained with the 12-in. stimulus as the CS, and each group was extinguished with a different conditioned test stimulus.

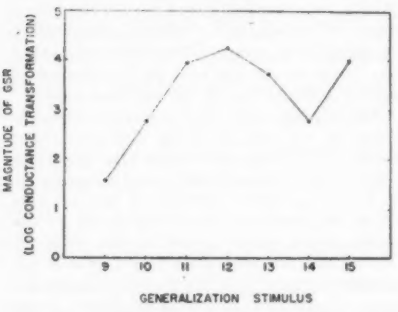


FIG. 1. Averaged GSR magnitudes in log conductance units to different generalization stimuli on the first extinction test trial. Each point based on initial responses of 14–16 *S*s. All *S*s had been trained with the 12-in. CS.

RESULTS

All records were read in terms of millimeters deflection and were transformed to ohms resistance change from a calibration curve. Resistance measures were then transformed into conductance measures, and the score that was used for final statistical analysis was the natural logarithm of 1,000 times the change in conductance plus

$$1 \cdot X = ln\ 1000 \left(\frac{1}{R - \Delta R} - \frac{1}{R} + 1 \right).$$

The distribution of these scores was unimodal and essentially normal aside from a secondary mode resulting from zero response. This corroborated previous observations on statistical transformations for the GSR (5, 6, 7, 8, 19).

The basic data of the experiment appear in Fig. 1 in which the average magnitude of the transformed GSR measures is plotted against the size of the generalization stimulus. The data upon which Fig. 1 is based came exclusively from the first extinction trial so that Fig. 1 presents a relatively

[2] Owing to some confusion between *E*s it was subsequently discovered that 16 *S*s had been extinguished to the 11-in. test stimulus and 14 had been extinguished to the 14-in. stimulus. The circumstances did not permit a correction of this defect which did not interfere with the subsequent statistical analysis.

pure generalization gradient. On the average, the maximum response occurred to the 12-in. test stimulus which was the CS used during training. The left-hand portion of the curve is a classically perfect generalization function up to the original CS, the 12-in. stimulus. The right-hand portion of the curve, obtained from the larger generalization stimuli, shows a marked inversion in that the average magnitude of response to the 15-in. generalization stimulus attains almost the same magnitude as the average response to the original CS. This spoils what would otherwise be a classical textbook picture of a generalization function.

The trend of the data in Fig. 1 was analyzed by means of orthogonal polynomials (**3**, pp. 20–22). The seven points in Fig. 1 yield 6 *df* which can be separated into the six orthogonal components owing to linear trend, quadratic trend, cubic trend, etc. The summary of the analysis of variance appears in Table 1. Table 1 contains the mean squares and *F*'s with which to evaluate the significance of each of the six components of trend; the mean square for error is based on 98 *df*. The *F*'s of Table 1 show that there is a linear trend that is significant at the 5% level and that the quadratic trend is also significant at the 5% level, but that none of the higher polynomials

TABLE 1

ANALYSIS OF THE GENERALIZATION FUNCTION OF THE GSR BY MEANS OF ORTHOGONAL POLYNOMIALS

Source of Variation	*df*	Mean Square	*F*
1st degree component	1	26.58	4.37*
2nd degree component	1	25.68	4.22*
3rd degree component	1	16.36	2.69
4th degree component	1	11.07	1.82
5th degree component	1	0.26	0.04
6th degree component	1	1.18	0.19
Error	98	6.08	

* Significant at the 5% confidence level.

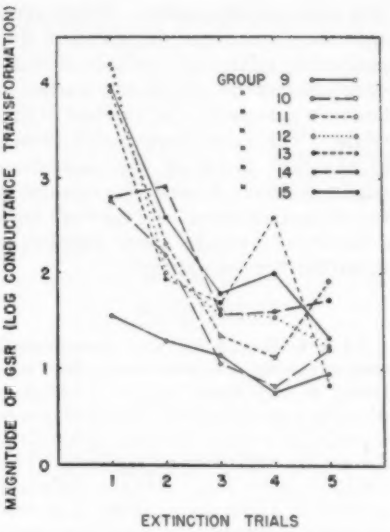


FIG. 2. Extinction curves for generalized GSR's on the first five unreinforced trials. Average response in log conductance units is plotted against extinction trial with the value of the generalization test stimulus as the parameter.

reveals significant variation. The significance of the linear trend indicates that there is a general upward trend in the means as one goes from the smaller to the larger generalization stimuli. The significance of the quadratic trend means that superimposed upon the upward linear trend is a curvilinear trend rising from Stimulus 9 to a maximum at Stimulus 12 and then dropping thereafter with the larger generalization stimuli.

Figure 2 contains the data from the first five extinction trials. The magnitude of the GSR on the first extinction trial gives the data already presented, and additional data are given for Trials 2, 3, 4, and 5. It will be noted that a more or less regular extinction trend appears for all groups. Furthermore, the orderliness among the seven generalization test groups which appeared in Fig. 1 has pretty well disappeared on the second extinc-

tion trial and thereafter. When averages were computed for the first five extinction trials, an analysis of variance showed no significant variation between groups for the average of five trials. The same irregularity shown in the first block of five extinction trials continues through the remaining ten extinction trials. At no time does a five-trial average show significant variation between groups.

Discussion

As is indicated by the experimental design, the main experimental effort was aimed at obtaining evidence regarding stimulus generalization. The design was kept simple in order that any generalization trends might become apparent. The results shown in Fig. 1 with the statistical analysis in Table 1 give fairly clear evidence for a statistically significant pure generalization curve. Unfortunately, superimposed upon this generalization curve was an upward trend in the data as the generalization stimuli increased in size. This may have been an effect caused by the magnitude of the test stimuli (13), either in length or possibly because of area-intensity summation (1). The fact that a very large number of degrees of freedom were involved and that the generalization effect was significant only at the 5% level points to the evanescent character of this phenomenon. Such an evanescent character would almost be postulated in view of the difficulties that previous *Es* have experienced in searching for such gradients.

The complexity of the extinction curves and the rapid loss of any clear-cut generalization trend in the extinction trials subsequent to the first as shown in Fig. 2 may be another reason why unambiguous demonstration of generalization and discrimination curves with the GSR has eluded so many investigators. If the data of the present experiment had been reported in terms of the average response for the first five trials, no clear-cut trend would have appeared. It would seem, therefore, that the pure generalization function itself is perhaps more reliable and clear-cut than the combined generalization and discrimination functions which develop after the first extinction trials.

Because of the simplicity of the experimental design it remains possible that phenomena other than conditioning are represented by the curves of Fig. 1 and 2. It is believed, however, that these other phenomena are minimized. For example, no pseudoconditioning (9) control trials were run, but the significant quadratic component of Fig. 1 could scarcely have been produced by simple pseudoconditioning, and, furthermore, the marked extinction shown in Fig. 2 has rarely been found in pseudoconditioning experiments. The results of the experiment might also be interpreted in terms of some frame of reference effect. All groups were trained to the 12-in. stimulus and the initial extinction response might have been some simple sensitization phenomenon linked up with the relation of the particular generalization test stimulus to the training CS. A frame of reference effect cannot be eliminated from generalization and discrimination processes since generalization gradients depend on the existence of some sort of continuum. Without a continuum the phenomena would be more or less meaningless.

Until evidence to the contrary is presented the results of this experiment indicate that there are basic generalization gradients for the conditioned GSR and that they are probably convex upward.

Summary

The nature of a pure generalization gradient for the conditioned GSR was investigated using 105 human *Ss*. The CS during training and generalization test stimuli were vertical rectangles projected upon a screen in front of *S*. The UCS was an electric shock. After 25 conditioning trials with a 1 × 12-in. rectangle as CS, different groups of 14–16 *Ss* were given 15 extinction trials with test stimuli 1 in. wide and 9, 10, 11, 12, 13, 14, or 15 in. in height.

Statistical analysis of the data revealed the following findings:

1. On the first extinction trial a pure generalization curve was obtained which was convex upward from the CS of 12 in. to the 9-in. test stimulus, and convex upward to the 14-in. test stimulus, but the results for the 15-in. test stimulus showed an inversion in the function. A progressive upward trend in average magnitude of the GSR from the 9-in. stimulus to the 15-in.

stimulus was significant at the 5% level. Superimposed upon this was a decreasing generalization function on both sides of the 12-in. training stimulus, which was significant at the 5% level. The significant linear trend was attributed to a stimulus magnitude effect possibly due to an area-intensity summation. The significant quadratic component in the findings was attributed to the progressively decreasing contribution of generalization as test stimuli were more remote from the CS.

2. After the first extinction trial a general extinction trend was obtained for all seven test stimuli. The orderliness due to the generalization function which appeared on the first extinction trial disappeared during the course of subsequent extinction. Statistical analysis of *five trial averages* during extinction showed no significant group differences.

It was pointed out that the results were probably caused by stimulus generalization. However, with present methods the phenomenon is evanescent, and large numbers of *S*s are required to demonstrate it.

References

1. Brown, R. H., & Niven, J. I. The relation between the foveal intensity threshold and length of an illuminated slit. *J. exp. Psychol.*, 1944, **34**, 464–476.
2. Davis, R. C. *Ability in social and racial classes: some physiological correlates.* New York: Century, 1932.
3. Fisher, R. A., & Yates, F. *Statistical tables for biological, agricultural, and medical research.* London: Oliver and Boyd, 1943.
4. Greenwald, D. U. A modification of Greenwald's apparatus for measuring electro-dermal responses. *Amer. J. Psychol.*, 1936, **48**, 508–509.
5. Haggard, E. A. Experimental studies in affective processes: II. On the quantification and evaluation of "measured" changes in skin resistance. *J. exp. Psychol.*, 1945, **35**, 46–56.
6. Haggard, E. A. On the application of analysis of variance to GSR data: I. The selection of an appropriate measure. *J. exp. Psychol.*, 1949, **39**, 378–392.
7. Haggard, E. A. On the application of analysis of variance to GSR data: II. Some effects of the use of inappropriate measures. *J. exp. Psychol.*, 1949, **39**, 861–867.
8. Haggard, E. A., & Garner, W. R. An empirical test of a derived measure of changes in skin resistance. *J. exp. Psychol.*, 1946, **36**, 59–70.
9. Harris, J. D. Studies on nonassociative factors inherent in conditioning. *Comp. Psychol. Monogr.*, 1943, **18**, Whole No. 93.
10. Hilgard, E. R. *Theories of learning.* New York: Appleton-Century, 1948.
11. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: Appleton-Century, 1940.
12. Hovland, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. gen. Psychol.*, 1937, **17**, 125–148.
13. Hovland, C. I. The generalization of conditioned responses: II. The sensory generalization of conditioned responses with varying intensities of tone. *J. genet. Psychol.*, 1937, **51**, 279–291.
14. Hovland, C. I. The generalization of conditioned responses: III. Extinction, spontaneous recovery, and disinhibition of conditioned and of generalized responses. *J. exp. Psychol.*, 1937, **21**, 47–62.
15. Hovland, C. I. The generalization of conditioned responses: IV. The effects of varying amounts of reinforcement upon the degree of generalization of conditioned responses. *J. exp. Psychol.*, 1937, **21**, 261–276.
16. Hovland, C. I. Comments on Littman's "Conditioned generalization of the galvanic skin reaction to tones." *J. exp. Psychol.*, 1949, **39**, 892–896.
17. Hull, C. L. The problem of primary stimulus generalization. *Psychol. Rev.*, 1947, **54**, 120–134.
18. Humphreys, L. G. Generalization as a function of the method of reinforcement. *J. exp. Psychol.*, 1939, **25**, 361–372.
19. Lacey, O. L., & Siegel, P. S. An analysis of the unit of measurement of the galvanic skin response. *J. exp. Psychol.*, 1949, **39**, 122–127.
20. Lashley, K. S., & Wade, M. The Pavlovian theory of generalization. *Psychol. Rev.*, 1946, **53**, 72–87.
21. Littman, R. A. Conditioned generalization of the galvanic skin reaction to tones. *J. exp. Psychol.*, 1949, **39**, 868–882.
22. Muenzinger, K. F., & Walz, F. C. An examination of electrical-current-stabilizing devices for psychological experiments. *J. gen. Psychol.*, 1934, **10**, 477–482.
23. Razran, G. Stimulus generalization of conditioned responses. *Psychol. Bull.*, 1949, **46**, 337–365.
24. Wickens, D. D. Stimulus generalization of the GSR. Paper read at Midwest. Psychol. Ass., Chicago, 1951. (Abstract, ADI Document No. 3277)

(Received January 29, 1953)

Journal of Experimental Psychology
Vol. 46, No. 5, 1953

SCOTOPIC AREA–INTENSITY RELATIONS AT VARIOUS RETINAL LOCATIONS [1]

ARTHUR J. RIOPELLE

Emory University

AND KAO LIANG CHOW

Yerkes Laboratories of Primate Biology

All descriptions of the area-intensity relation have shown that an increase in the size of a visual stimulus results in decreased threshold. The most recent determinations of the relationship between these variables have been presented by Wald (12) and by Graham, Brown, and Mote (6). These authors have also provided theoretical formulations to account for the empirical results. The purpose of the present experiment was to determine whether the area-intensity relation changed as a function of changes in the location of stimulation. Such a determination would provide information as to the generality of the proposed theoretical formulations and also would be useful for revealing the significance of regional variations in particular characteristics of retinal structures.

METHOD

Apparatus.—Light from an automobile headlamp was directed through a Wratten neutral-tint wedge and balancer, fixed filters, lenses, and a ground-glass diffuser to fall eventually upon an opal glass screen. The reverse side of this screen, when viewed through a hole drilled in an opaque mask, served as the stimulus. Brightness of the screen was determined photometrically with the aid of a Macbeth illuminometer. Ten masks were used, the holes in which varied in diameter from 1/16 in. to ½ in. When viewed at a distance of 30.5 in. the holes subtended visual angles ranging from 7 min. to 57 min. The largest area was therefore 64 times greater than the smallest. Appropriately placed fixation lights provided stimulation of retinal regions located at 4, 8, 12, 24, or 48 degrees above the fovea.

At one point the beam of light passed through a lens which collected the rays. The passage of an open sector in a rotating disk through the focal point of this lens permitted a brief pulse of light to reach the stimulus screen.[2] The rise and fall times of the pulse were small portions of the total duration.

Subjects.—The two authors alternated as *S* and as *E*. Both *S*s had undergone several practice sessions prior to the experiment proper and both were well acquainted with their tasks. In addition, a portion of the time at the beginning of each experimental session spent in the preliminary dark adapting was devoted to practice, thereby enabling *S* to define a stable criterion of visibility.

Procedure.—The *S* was seated in a light-tight cubicle with his head held immobile by a chin and forehead rest. After 30 min. of dark adaptation *S* turned his gaze to a small fixation light which was located to provide stimulation of the desired retinal location. The brightness of this fixation light was controlled by *S*, and was kept as dim as possible consistent with effortless fixation.

Thresholds were taken by a modified method of limits. After two dim but definitely supraliminal flashes were presented, the intensity of the flash was reduced to a value below threshold. Subsequent flashes increased in intensity by discrete steps of .01–.05 log micromicrolamberts until *S* reported seeing two successive flashes. Five such thresholds were obtained in succession for a particular size of stimulus. The median thresholds in terms of log micromicrolamberts were the experimental values used in the analysis.

Two flash durations were used: 50 and 250 msec. These values were chosen as representative of durations shorter and longer than the critical duration (5), the point at which the duration of the flash ceases to be a variable in the determination of the absolute threshold. The interflash interval was always greater than one second.

Two area-intensity curves were completed in

[1] Supported in part by a grant from the University Center in Georgia, Inc.

[2] Thanks are due Mr. R. A. Hall and his associates at the Experiment Station of the Georgia School of Technology for designing and constructing the shutter mechanism.

a single experimental session. One of these was based on flash durations of 50 msec., the other of 250 msec. Both curves were obtained at the same retinal location. Since dark adaptation might be expected to continue beyond the preliminary 30 min., the experimental conditions were balanced with respect to duration of dark adaptation. Because of this balancing any effect of differential dark adaptation can be isolated in the statistical analysis.

Results

In Table 1 are shown the mean thresholds obtained at the various retinal locations. These data show agreement with previously published gradients of sensitivity for various eccentricities of stimulation (**4, 9, 10**)[3] in that the point of maximum sensitivity lies some distance from the fovea.

For both *S*s the mean threshold differences for the two flash durations were in the neighborhood of .47 log units; the longer the flash duration, the lower the threshold intensity. Individual differences in over-all sensitivity amounted to .1 log unit, those for KLC being consistently the lower. The thresholds taken during the first run of a session, i.e., after only 30-min. dark adaptation, were for AJR the same as those taken during the second run (after 70-min. dark adaptation). For KLC, however, the thresholds of the second run (after 90 min.) were slightly higher. These data are presented in Table 2. Another effect which was shown in the data of KLC, but not of AJR, also related to differential dark adaptation. Extreme peripheral thresholds obtained after 90-min. dark adaptation tended to be higher for this *S* than those obtained after only 30-min. dark adaptation. The observed effect in both instances is probably best attributed to an irrelevant variable, possibly fatigue. The effects of the various factors just mentioned have been analyzed for two reasons: first, to corroborate previous data, and second, to eliminate the effects of known sources of variation during consideration of the principal variable.

TABLE 1

Mean Thresholds in Log Micromicrolamberts at Various Retinal Locations

Location	Eccentricity in Degrees	Subjects	
		KLC	AJR
1	4	5.65	5.76
2	8	5.56	5.70
3	12	5.60	5.70
4	24	5.73	5.82
5	48	5.79	5.81

TABLE 2

Mean Thresholds after Varying Amounts of Dark Adaptation

Subject KLC		Subject AJR	
Time	Log μμl	Time	Log μμl
30 Min.	5.63	30 Min.	5.76
90 Min.	5.70	70 Min.	5.75

The phenomenon of primary interest in this experiment involves the interrelation of retinal position and the area-intensity effect. The data bearing on this are derived from a comparison of the area-intensity curves obtained at the several retinal locations. These results are portrayed graphically in the two panels of Fig. 1. This figure shows area-intensity relations for the two *S*s at each retinal position. The data upon which these curves are based are derived from averages over both flash intervals and both durations of dark adaptation. Each curve is thus a composite of four curves. The notations of the curves denote the particular retinal location where the area-intensity relation was determined.

The striking feature of these curves

[3] Sloan, L. L. Personal communication.

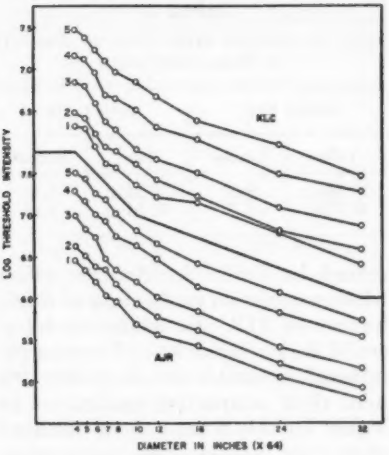


FIG. 1. Threshold intensity as a function of the diameter and the retinal location of the test patch. The test patch was viewed at a distance of 30.5 in. Ordinate values apply only to the bottom curve of each panel, all other curves being displaced vertically successive multiples of .1 log micromicrolambert. The curve numbers refer to differing distances from the fovea; the higher the number the greater the distance.

is their parallelism. Although there may be slight differences in the constants of the curves between the two *S*s, it is clear that for a given *S* one curve may be made to coincide with another merely by adding appropriate constants to the ordinate values of the curves.

Separate analyses of variance were performed on the data of the two *S*s. These analyses are summarized in Table 3. Inspection of this table reveals that the results cited above are supported by the statistical analysis. Three of the main effects are significant for both *S*s, and a fourth main effect, attributable to differential dark adaptation, is significant for KLC. A single significant first-order interaction was also found in the data of that *S*. The sums of squares and degrees of freedom of all higher-order interactions were pooled to provide a stable error estimate.

It was concluded that the curves of Fig. 1 were parallel. This conclusion is supported by the nonsignificant *F*'s for the interaction of *Location by Size*. Both *F*'s were less than 1.00, and the precision of the experiment, as judged by the smallness of the error mean square and the many degrees of freedom, is such as to have provided ample opportunity for such an effect to be detected.

DISCUSSION

This experiment was designed to determine whether any changes occur in the area-intensity relation with changes in the locus of retinal stimulation. We have concluded that for the range of areas and the flash durations used, no such changes occur. Thus spatial summation within an area of homogeneous stimulation is equally efficient at markedly different distances from the fovea.

There are probably severe restrictions upon the generality of the results. First, the areas of stimulation employed were relatively small, the largest subtending less than one degree of visual angle. The use of stimuli larger than one degree might account for Wald's discovery of separate area-intensity curves at two dif-

TABLE 3

SUMMARY OF ANALYSES OF VARIANCE

Source of Variation	*df*	Mean Square	
		KLC	AJR
Retinal location (L)	4	.356**	.140**
Flash duration (D)	1	10.452**	10.695**
Dark adaptation (DA)	1	.220**	.004
Stimulus area (S)	9	6.235**	6.620**
L × D	4	.020	.013
L × DA	4	.075**	.011
L × S	36	.008	.008
D × DA	1	.017	.011
D × S	9	.005	.031
DA × S	9	.002	.004
Higher interactions	121	.012	.015
Total	199		

** Significant at 1% level of confidence.

ferent retinal locations (12). His smallest area subtended one degree of visual angle. Second, the deterioration of the linear reciprocity of area, intensity, and duration for large areas and for long durations (5) leads one to expect that the additive effect of flash duration is restricted to small areas of stimulation.

The finding of nonfortuitous parallelism in the curves at five different locations bears upon current conceptions as to the mechanism underlying the reciprocity of area and intensity. The general theoretical account advanced by Wald (12) proposes a statistical distribution of the thresholds of receptive units. Increasing the stimulus area causes stimulation of a greater number of low threshold units. Specification of the frequency distribution of the thresholds involves both a measure of over-all level of sensitivity and a measure of the variability of the receptor population. In terms of Wald's hypothesis the present data imply that although the general level of sensitivity may change somewhat along a meridian, the variability in receptor thresholds within a small area remains constant. The implication of these results for Crozier's theory of the statistical distribution of neural events would be unchanged in its essentials (3).

The theory proposed by Graham, Brown, and Mote (6) to account for the area-intensity effect emphasizes the interaction between different parts of the stimulated area. Their formulation states that the contribution of an elementary area to the total excitatory effect is inversely proportional to a power, p, of its distance from the center of stimulation. The present results imply that the summation factor, p, remains constant throughout the peripheral retina. This may, in turn, imply some uniformity in the underlying mechanism.

Various attempts to relate spatial summation to the anatomy of the visual system have emphasized retinal structures, particularly after Adrian and Matthews' demonstration (1) of summation in the excised eel's eye. Of special interest is the interrelation of visual summative phenomena with the spatial distribution of receptors, bipolars, and ganglion cells. One example of this interest is the comparison (4, 9) of the gradient of threshold sensitivity with the density gradient of rods as determined by Østerberg (8).

The greater density of rods in the peripheral retina in comparison with the density of bipolar and ganglion cells has led to the interpretation of "convergence" of retinal effect. Bartley (2, p. 86), for example, has stated that 80 rods may converge upon each optic nerve fiber. The ratio of rods to bipolars has been believed to vary with increasing distance from the fovea to the periphery, even attaining an estimated maximum of 200:1 (7, p. 164). A gradient of convergence across the retina would imply differential efficiency in spatial summation. Vilter (11), however, has counted the retinal elements and has found that practically everywhere in the retina except the fovea 100 rods converge to 17 bipolars which in turn impinge upon one ganglion cell. He found no appreciable variation in the degree of convergence. It is probably significant to note that this uniformity of convergence of structural elements is correspondent with a uniformity in the efficiency of spatial summation, at least as measured by the area-intensity effect.

Summary

Absolute scotopic thresholds were determined for ten areas of stimulation at each of five different retinal locations. All thresholds were made monocularly with the naked eye following at least 30 min. of dark adaptation. The largest area (subtending 57 min. of visual angle) was 64 times larger than the smallest area. The locations of retinal stimulation were 4, 8, 12, 24, and 48 degrees above the fovea. Stimulus durations were 50 and 250 msec. The authors alternated as *S* and as *E*. The results may be summarized as follows: (*a*) There was a significant difference in over-all threshold at the various retinal locations. (*b*) The longer the duration of the test flash, the lower the intensity required for threshold. (*c*) The larger the stimulus patch, the lower the threshold. (*d*) Area-intensity curves obtained at different retinal locations were parallel, indicating that within the range of this experiment the area-intensity relation is invariant with respect to locus of its determination.

References

1. Adrian, E. D., & Matthews, R. The action of light on the eye. III. The interaction of retinal neurones. *J. Physiol.*, 1928, **65**, 273–298.
2. Bartley, S. H. *Vision, a study of its basis.* New York: Van Nostrand, 1941.
3. Crozier, W. J. On the sensory discrimination of intensities. *Proc. nat. Acad. Sci.*, 1936, **22**, 412–416.
4. Crozier, W. J. Theory and measurement of visual mechanisms: I. A visual discriminometer. II. Threshold stimulus intensity and retinal position. *J. gen. Physiol.*, 1939, **22**, 341–364.
5. Graham, C. H., & Margaria, R. Area and the intensity-time relation in the peripheral retina. *Amer. J. Physiol.*, 1935, **113**, 299–305.
6. Graham, C. H., Brown, R. H., & Mote, F. A., Jr. The relation of size of stimulus and intensity in the human eye: I. Intensity thresholds for white light. *J. exp. Psychol.*, 1939, **24**, 555–573.
7. Morgan, C. T. *Physiological psychology*, New York: McGraw-Hill, 1943.
8. Østerberg, G. Topography of the layer of rods and cones in the human retina. *Acta Opthal. Suppl.*, 1935, **61**, 1–102.
9. Riopelle, A. J., & Bevan, W. The distribution of scotopic sensitivity in human vision. *Amer. J. Psychol.*, 1953, **66**, 73–80.
10. Sloan, L. L. Rate of dark adaptation and regional threshold gradient of the dark-adapted eye: physiologic and clinical studies. *Amer. J. Ophthal.*, 1947, **30**, 705–720.
11. Vilter, V. Recherches biométriques sur l'organisation synaptique de la rétine humaine. *C. R. Soc. Biol.*, 1949, **143**, 830–832.
12. Wald, G. Area and visual threshold. *J. gen. Physiol.*, 1938, **21**, 269–287.

(Received February 28, 1953)

Journal of Experimental Psychology
Vol. 46, No. 5, 1953

CHARACTERISTICS OF THE MUSCLE TENSION RESPONSE TO PAIRED TONES[1]

D. W. VAN LIERE[2]
Indiana University

It has been found with the various sense modalities and under varying experimental conditions that the repeated presentation of a given stimulus generally alters in some way the manner in which *S* responds to that stimulus, whether *S* be animal or human. Conditioning is a case in point. The manner in which the response undergoes change may be measured in terms of the latency of the response, the frequency of response, the amplitude of the response. Any one of these or any combination of them may undergo an evolutionary alteration. These changes have usually been studied in terms of an overt response. But the persistent question has been raised whether an overt response is necessary to conditioning. Association theory of the stimulus-stimulus variety suggests there need not be any kind of overt response when conditioning occurs (21).[3]

We seem to have a nonovert response in the muscle tension reflex induced by an auditory stimulus. Davis has determined that a measurable EMG can be produced with near-threshold stimuli, and in his discussion of results advances as one of his explanatory principles that of "partial forward displacement" which, in his opinion, may resemble an incomplete CR (7). The present study then is concerned with investigating the muscle tension reflex in a typical classical conditioning paradigm to see whether a CR can be established where no "overt" response is involved.

Apparently only three direct attempts have been made to deal with the conditioning of the EMG (12, 16, 17). In the first of these studies, with EMG's in the white rat, Prosser and Hunter conclude that the results of the process of pairing a subthreshold sound and a tetanic shock were not indicative of true conditioning but "merely a raising of the excitability of the centers so that a sub-threshold stimulus becomes and then remains effective" (17, p. 613). Using a light of .7 ft.-candles as CS and an electric shock as US, Hunter (16) obtained a well-established CR where no response to the light was initially present. By use of needle electrodes he was able to show that the conditioning process brought about the firing of individual units not before elicited by the light.

Hilden (12) has studied action currents in a conditioned hand withdrawal. Using a light-shock method he established an EMG-CR with a latency about the same as for a voluntary response. Further, he found that the EMG–CR lasts very much longer than the overtly measured CR. This suggests that an EMG could perhaps be conditioned without the presence of an overt response.

The present study is a re-examination of the conditioning of the muscle tension reflex under somewhat different conditions: (*a*) no overt response to the US is involved; (*b*) the stimuli for the US and CS are mediated by the same sense modality. The classical conditioning paradigm is used, with CS and US differing with respect to the physical properties of amplitude and frequency of the stimuli. We are concerned with the process or proc-

[1] This paper is based on a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, in the Department of Psychology, Indiana University, 1950. The writer is indebted to Dr. R. C. Davis for his suggestions, criticisms, and encouragement in completing this research and for the use of his laboratory in collecting the data.

[2] Now at Kalamazoo College, Kalamazoo, Michigan.

[3] "Overt" is relative to the measuring instrument used. Here we refer to it in terms of finger withdrawal or some such similar movement of a body measurement through a measurable distance in space.

esses which occur when these stimuli are paired for 25 trials. If we have a simple associative, noninstrumental conditioning process, then we should expect an increase in the size of the response to that which, in conditioning terminology, is the CS. In a similar manner, we may say that whenever a response is elicited after pairing which was not elicited before such pairing then we have evidence of a conditioning process (13).

Method

Apparatus.—The apparatus and techniques of recording and electrode placement were similar to those reported elsewhere by Davis (2, 3, 4, 6, 7). The stimuli for the CS and US were produced by two Hewlett-Packard audio oscillators; a CS of 850 cps and a UCS of 1,000 cps delivered through a loud-speaker. Amplitude control was effected at the oscillators through a General Radio Audio-microvolter and by the electronic timer modified from Roush (19). Intensities were determined with a General Radio Sound Level Meter. With the pick-up in the position of *S*'s head, the CS was 43 db and the UCS was 70 db. The electronic timer produced the CS for 1.77 ± .010 sec. and the UCS for 2.00 ± .024 sec. The UCS began about .04 sec. before the CS was terminated. The sequence was initiated by *E*'s activating the timer which was synchronized with the integrator system.

Electrodes were placed on right and left masseters and on the extensor muscles on the dorsal surface of the right and left forearms. Reference electrodes on the upper arm or ear lobe follow Davis' technique.

Subjects.—The *S*s for the present experiment were 48 (24 experimental and 24 control) volunteer, unselected students from introductory experimental psychology classes.

Procedure.—The *S*s lay relaxed in a shielded, dark, sound-insulated room. Each *S* received the following instructions: "In this experiment we are studying the effects of sounds. From time to time you will hear tones. There are no responses to make to these tones. You will merely be as relaxed as possible at all times." No interruptions were introduced once the trials were begun. Trials were given on an average of about one per minute. Intertrial interval was randomly varied to prevent possible response to a temporal interval. Twenty-five trials were presented to each group. The control group received only the CS of 850 cps and 43 db. The experimental group received the paired tones on all except the test trials (Trials 1, 5, 10, 16, 21–25, the last five constituting an "extinction" series). The effect of the pairing could thus be determined by comparing the responses of the control group and the experimental group on Trials 1, 5, 10, 16, and 21–25.

Measurement of response.—Responses were summated over a time interval of 7/60 sec. by means of an electrical integrator. Thus individual spikes were added for 7/60 sec.; the integrators were then shorted and reset to begin summation of spikes again. These summations are referred to as integrations and are counted from the point of stimulus onset.

A plot of the mean of all muscles and all trials showed an *a* response to appear in the vicinity of integrations 3, 4, 5, and 6 following stimulus onset, and a *b* response for integrations 11, 12, 13, and 14 (5). For both the experimental and control groups this would be the response to the CS. Measuring the response to the UCS was also made from the onset of the US. (We will designate the responses to the UCS as *a'* and *b'* responses to differentiate them from the *a* and *b* responses to the CS.) Since the control group received no UCS the equivalent of the *a'* and *b'* responses were measured from the point of cessation of the CS.

Response values are in terms of the percentage increase or decrease of tension over a prestimulus tension level (2). This means that for each trial *S*'s prestimulus level of muscle tension was determined and the response to the stimulus was calculated as a percentage of the prestimulus level. Thus, suppose that *S* showed a prestimulus tension level of 10 microvolts in the right masseter and following the stimulus *S* gave a 15-microvolt *a* response. This, when transposed into percentage of the prestimulus level, is a 150% response. The prestimulus level was determined by taking the highest integration from integrations 5, 4, 3, and 2 preceding the stimulus onset, and the *a* response level was determined by taking the highest integration from integrations 3, 4, 5, and 6 following stimulus onset (6).

Results

The presence of responses.—It becomes necessary to establish the existence of the responses with which we are to be concerned. In all cases we are comparing the obtained response amplitude with a hypothetical value. If there were no response, then the obtained mean response amplitude would not differ significantly from zero. Statistically we assume a null

hypothesis which states that the difference between the obtained and hypothetical mean can be accounted for on the basis of chance factors alone.

On the first trial neither experimental nor control group showed a significant response to the 850-cps, 43-db. tone (CS). By Trial 5, however, the control group gave significant t values for the a and b responses ($p<.01$), and the experimental group gave significant t's for the a response ($p<.02$) and for the b response ($p<.05$). In the experimental group the responses to the UCS, the a' and b' responses, were significant beyond the 1% level. We may thus conclude that our *S*s did make muscular responses to the stimuli used.

Responses in the experimental and control groups.—On the initial trial for the experimental and control groups we find that there was no significant difference between them as measured by the t test of the mean a responses (Fig. 1). There is also no significant difference between the mean b responses for the two groups on the initial trial (see upper curves in Fig. 1). Analysis of variance for a and b responses from all locations was used to test this difference in trend. The question is whether the apparent group difference is attributable to within-group individual variations in trend. Here the trial × group interaction was tested by the pooled trial × individual interaction within groups (10). Comparing the experimental and control groups we get an F of 5.64, which is significant between the .05 and .01 levels of confidence. It appears, therefore, that the pairing of the CS and UCS had an effect upon the response to the CS, and that a process of conditioning in skeletal muscle was operating.

An examination of the data for specific response or muscle position differences by means of the t test (with the error term being simply individual differences) reveals that the a response was significantly greater in the experimental group ($t = 2.12$; $.05>p>.02$). A similar test for the b response gives a nonsignificant t of 1.79. This is not very meaningful since the over-all mean difference is diluted by the initial near-identity of the responses on Trial 1.

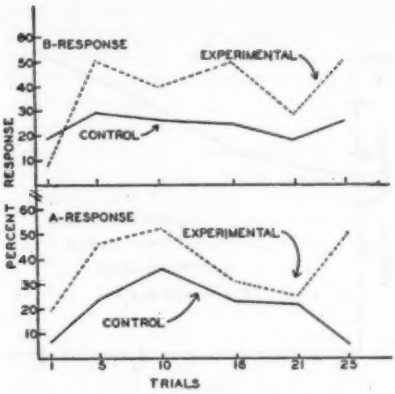


FIG. 1. The amplitude of a and b responses for experimental and control groups on test trials

A comparison of the muscle positions shows that no one muscle is more active than another. Although in our experimental group the masseters appear to be more active, this does not prove to be significantly so.

The conclusion must be that there is a significant response difference in our experimental group and that it is to be found primarily in the a response and is not specific to any of the muscle groups examined. Skeletal muscle appears to be conditionable as measured by the EMG although it cannot be indicated that any one of the particular muscles measured in the present study is thus conditionable.

Effect of the pairing on the UCR.—The present study shows that the CS

FIG. 2. Increased UCR as a result of pairing with the CS

has a facilitating effect on the UCR ($p < .05$). By comparing four successive means of groups of three trials each we note the increasing trend (Fig. 2). An analysis of variance applied to these data shows that the increase from 35.23% response to 60.58% response is significant beyond the 5% level ($F = 4.31$).

This phenomenon has been reported previously by Schlosberg (20). He found a high correlation between the success of conditioning and the amount of facilitation of the UCR by the CS. Razran found a similar phenomenon in salivary conditioning in humans (18). Harris reports that the size of the UCR increases as buzz-shock conditioning occurs (11). Hilgard has shown, however, that at the favorable intervals for eyelid conditioning the CS definitely depresses the response to the UCS (13).

Supplementary findings.—In addition to the above there are a number of other findings of interest in connection with a conditioning situation: (*a*) There is evidence of an extinction phenomenon similar to that first reported by Switzer (22) and later by Hudgins (15) and Hovland (14). Switzer (22) reported a novel form of experimental extinction curve. The amplitude of a conditioned eyelid reaction was found to rise during the initial trials of extinction so that Trials 2 and 3 were larger than was the initial extinction trial. In the present study a similar extinction curve is observed and a statistical comparison of Trials 1 and 3 of the extinction series reveals a difference which is significant at the 2% level. This increase in response during the early trials of extinction is still reflected in the fifth extinction trial (see Trials 21 and 25 in Fig. 1). (*b*) There is a sensitization-adaptation process in the control group differing from that previously reported by Davis (2, 5, 6) and Davis and Van Liere (9), but in agreement with Davis' (7) report of a sensitization to weak stimuli. In the present study an analysis of variance shows a significant trial-to-trial change beyond the 1% level. A breakdown by *t* test shows Trial 5 with an *a* response significantly greater ($p < .01$) than that of Trial 1. Trial 10 shows an *a* response four times the size of the initial trial, and on Trial 25 the response is as strongly present as on the initial trial.

DISCUSSION

An examination of the studies of the muscle tension response produced by sound shows that each succeeding study gives a somewhat different picture of the adaptation phenomenon. Differences in the frequency of the stimulating sound, differences in intensities, differences in *S*'s activity all seem to complicate the formulation of an hypothesis concerning the adaptation phenomenon. In dealing with stimuli at the absolute limen Davis (7) found significant increases in the size of the *a* and *b* responses as the trials progressed. In interpreting his results he advances the principle of forward displacement which he says "sounds a great deal like some accounts of the conditioned response. What occurs here may indeed be a conditioned response in a very incomplete form" (7, p. 118). The present study attempts to answer this by use of the conditioning paradigm. Results show an increase in the size of both

a and *b* responses to the CS. However, this is true of both experimental and control groups. And in the experimental group it is true of both the CR and UCR. It would seem then that the sensitization process is not merely a conditioning process. If we consider the CR alone (the response to the weaker tone), we may say that sensitization may occur with weak stimuli while adaptation occurs with strong auditory stimuli. However, a similar process occurs with the UCR (response to the stronger tone). This of course may be due to the facilitating effect of the CS on the UCR which was mentioned above. Since we had no group receiving the UCS only, we can add nothing in answer to the question of the relation of tone intensity to the sensitization process.

There are seemingly common factors operating in both experimental and control groups. The factor of the weaker tone is not an adequately explanatory one. A possible explanation for our sensitization might be in an increased level of tension in the experimental situation as a whole; that is, the lack of adaptation may be a function of a generally increased level of tension in *S*. Evidence seems to indicate that this is not the case. We find that there is an adaptation of the general tension level (as measured by the prestimulus tension level) during the early trials while the response level is going up. From Trial 5 on, the level of prestimulus tension in microvolts remained essentially constant in both experimental and control groups.

The factor of conditioning must not be overlooked, however. Since recent findings for the GSR in a conditioning situation show a similar sensitization followed by adaptation (1), we may not need to look to the nature of the stimulus or the experimental situation for an explanation. It may be something inherent in the process of conditioning. Conditioning may be a kind of sensitization process. In conditioning we may have an "expectation" which exaggerates response trends already inherent in the organism. The present study seems to bear this out in that the response to a weak tone is facilitated when it is followed by a stronger tone. In our case the response is initially present (although not significantly so) and so would not seem to be in conformity to the usual conception of the conditioning process.

Yet if we conceive of a CR to be dependent on the potentiality of being elicited with pairing, then may we not conclude that this state is actually present but not measurable by certain measuring devices? The salivary response in a dog is not elicited by a bell prior to pairing with food powder—but perhaps only in terms of a measurable quantity of saliva. A more refined measuring device might measure a change in the salivary gland which is enhanced or facilitated by pairing with powder so that salivation does occur. Perhaps the bell does produce a "neuromuscular" change in the salivary gland but one that is insufficient to produce salivation. The process of conditioning merely raises the "neuromuscular" change above a threshold, which threshold is determined as a measurable quantity of saliva.

The present study suggests just such a process. By working below the threshold of an overt response we are dealing with a measurable "potentiality." This may give the appearance of a response already present before conditioning. However, conditioning may merely be the process of raising certain neuromuscular reactions above a certain threshold so that a previously defined response can be measurably elicited. The present study indicates that a process of sensitization is enhanced by pairing with a stronger tone. It conforms to Hunter's (16) finding that in conditioning there is an increase in the number of active units (integrations of responses from surface electrodes would appear as greater total responses in the present study). Thus the bell as CS in the salivary conditioning experiment may activate an insufficient number of units to produce a measurable quantity of saliva. The pairing process (conditioning) increases the number of active units and hence salivation occurs.

With respect to conditioning, then, the present study seems to show: (*a*) that the pairings of responses to auditory stimuli simply exaggerate the trends of responses

that are already present; (*b*) that in one sense, at least, we have conditioning; (*c*) that such an effect may reconcile so-called sensory conditioning with response conditioning; (*d*) that simple contiguity seems adequate to produce conditioning of the EMG of skeletal muscles; and (*e*) that the sensitization phase may be the basis for pseudoconditioning.

Summary

This study was concerned with the conditioning and adaptation of the muscle tension reflex. With an 850-cps, 43-db tone as the CS and a 1,000-cps, 70-db tone as the UCS, the CS–UCS pair was presented for a total of 25 trials (with interspersed test trials, and five terminal "extinction" trials) to an experimental group ($N = 24$) and the CS was presented alone for 25 trials to a control group ($N = 24$). Electrical measurements of the muscle tension response to these tones were obtained from electrodes placed on the right and left masseters and on the extensor muscles on the dorsal surfaces of the right and left forearms.

Although the two groups did not differ initially in amplitude of response to the CS, both groups showed a significant increase (sensitization) in that response during Trials 1–5. The experimental group showed a significantly greater amplitude of response to the CS than did the control group. Also, there was evidence of a facilitative effect of the CS on the amplitude of the response to the UCS. These results are discussed in terms of their significance for conditioning.

References

1. Bitterman, M. E., & Holtzman, W. H. Conditioning and extinction of galvanic skin response in relation to clinical evidence of anxiety. Report No. 3, Project 21-37-002, School of Aviation Med., Randolph AFB, Texas, 1952.
2. Davis, R. C. Methods of measuring muscular tension. *Psychol. Bull.*, 1942, **39**, 329–346.
3. Davis, R. C. The design and testing of multiple amplifiers for action potential recording. *J. exp. Psychol.*, 1943, **32**, 270–274.
4. Davis, R. C. An integrator and accessory apparatus for recording action potentials. *Amer. J. Psychol.*, 1948, **61**, 100–104.
5. Davis, R. C. Motor effects of strong auditory stimuli. *J. exp. Psychol.*, 1948, **38**, 257–275.
6. Davis, R. C. Responses to 'meaningful' and 'meaningless' sounds. *J. exp. Psychol.*, 1948, **38**, 744–756.
7. Davis, R. C. Motor responses to auditory stimuli above and below threshold. *J. exp. Psychol.*, 1950, **40**, 107–120.
8. Davis, R. C. The stimulus trace in effectors and its relation to judgment responses. *J. exp. Psychol.*, 1952, **44**, 377–390.
9. Davis, R. C., & Van Liere, D. W. Adaptation of the muscular tension response to gunfire. *J. exp. Psychol.*, 1949, **39**, 114–117.
10. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
11. Harris, J. D. Facilitation of the unconditioned responses by the conditioned stimulus in buzz-shock conditioning of rats. *Psychol. Bull.*, 1942, **39**, 598. (Abstract)
12. Hilden, A. H. An action current study of the conditioned hand withdrawal. In University of Iowa studies in psychology. XXII. Studies in clinical psychology. *Psychol. Monogr.*, 1937, **49**, No. 1 (Whole No. 217), 173–204.
13. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: D. Appleton-Century, 1940.
14. Hovland, C. I. "Inhibition of reinforcement" and phenomena of experimental extinction. *Proc. nat. Acad. Sci.*, 1936, **22**, 430–433.
15. Hudgins, C. V. Conditioning and the voluntary control of the pupillary light reflex. *J. gen. Psychol.*, 1933, **8**, 3–51.
16. Hunter, W. S. Muscle potentials and conditioning in the rat. *J. exp. Psychol.*, 1937, **21**, 611–624.
17. Prosser, C. L., & Hunter, W. S. The extinction of startle responses and spinal reflexes in the white rat. *Amer. J. Physiol.*, 1936, **117**, 609–618.
18. Razran, G. H. S. Conditioned responses: an experimental study and a theoretical analysis. *Arch. Psychol.*, 1935, No. 191.
19. Roush, R. G. Tone burst generator. *Electronics*, 1947, **20**, 92–96.
20. Schlosberg, H. An investigation of certain factors related to ease of conditioning. *J. gen. Psychol.*, 1932, **7**, 328–342.
21. Spence, K. W. Theoretical interpretations of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951.
22. Switzer, S. A. Backward conditioning of the lid reflex. *J. exp. Psychol.*, 1930, **13**, 76–97.

(Received February 4, 1953)

Journal of Experimental Psychology
Vol. 46, No. 5, 1953

PERCEPTUAL JUDGMENT AS A FUNCTION OF MENTAL SET, ANCHORING POINT, AND METHOD OF JUDGMENT [1]

CHARLES W. HILL

Department of the Army

The purpose of this study was to integrate two heretofore separate lines of investigation into the dimensions of perceptual judgment. Recent studies (**4, 6, 8**) of judgment in psychophysics have emphasized the concept of the "anchoring point." The evidence indicates that the anchoring point is a stabilizing factor of great assistance in the relating of perceptual judgments to their referent stimuli. Meanwhile, social psychologists have focused their attention upon "mental set" in relation to judgment (**1, 2, 3, 7**). These experiments indicate that a "false" mental set produced by "suggestion" or "prestige" can be a very distorting influence. There seem to have been no prior studies containing both of these variables in opposition.[2]

The present experiment was designed to investigate the interacting influences of a visual anchoring point and a conflicting mental set induced by suggestion upon the formation of a scale of perceptual judgment. A third independent variable was introduced by using two methods of judgment, total impression and parts analysis, for the secondary purpose of determining whether or not this factor would affect the primary interaction.

It was hypothesized that (*a*) the suggestion-induced set would influence the judgments, and (*b*) both the anchoring point and the analytic method of judgment would tend to counteract this distortion.

METHOD

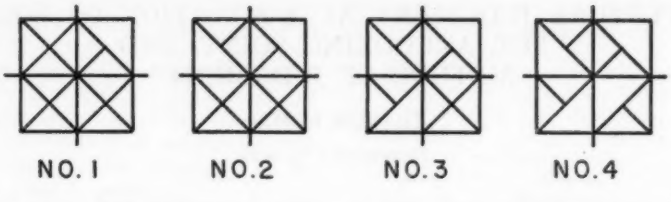
A series of unique, geometric figures was designed (see Fig. 1) in seven categories along the continuum of symmetry. As symmetry involves the correspondence of size, shape, and position of parts that are on opposite sides of a dividing line, the systematic variation in the lines of the stimulus figures about their central axes should produce systematic increases and decreases in symmetry. This a priori, objective scale was verified by ten judges sorting two sets of the figures into the seven categories according to the method of equal-appearing intervals.

Forty-six figures were then constructed with 2, 4, 10, 14, 10, 4, and 2 respectively in each of the seven categories. Within each category, a variety of figures was obtained by rotating the examples shown in Fig. 1 about the four quadrants. An additional figure of average symmetry, similar to Category 4 in Fig. 1, was constructed to serve as the anchoring point.

The stimulus figures were presented to 96 college students, one at a time, in a small, windowless room. The *E* sat behind a screen containing two small windows, one above the other, and exposed the figures in random order in the lower window for 5 sec. each by means of a shutter controlled by an electronic timer. A suggestion to the effect that most of the stimuli would be above average in symmetry was built into the instructions received by the "set" groups. The average figure was exposed in the upper window during the presentation of the series to the "anchoring-point" groups. The "total-impression" groups were told to use their over-all judgment, whereas the "parts-analysis" groups were encouraged to analyze the figures.

[1] This experiment was conducted at Vanderbilt University in partial fulfillment of the requirements for the Ph.D. degree. Dr. L. E. Thune deserves much credit and my deep appreciation for his assistance and advice, and my sincere thanks go to Dr. S. C. Ericksen for his helpful criticism.

[2] As Rogers pointed out in 1941 (**6**), suggestions and instructions as well as physical objects may be considered as anchoring points, and physical objects may also be thought of as suggestions to their perceiver. However, it is believed that the traditional usage of these terms is based upon meaningful distinctions and will be adhered to throughout this paper without further justification.



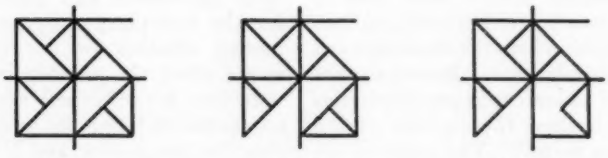


FIG. 1. Sample figures in the seven categories of symmetry

Seven-point graphic rating scales, with several phrases describing the various degrees of symmetry, provided a standard response form for the communication of *S*'s judgments. The phrases and direction of numbering were varied in a nonsystematic manner in an effort to disrupt the monotony of the rating process.

Eight experimental conditions were developed from the various combinations of the three dichotomized variables as shown in Table 1. As they arrived at the experimental room, *S*s were assigned to one of the conditions in accordance with a prearranged series of random numbers. Each of eight groups received 12 *S*s. The *E* verbally presented standardized instructions which varied among groups only as required by the variations in the three independent variables. During the 5-sec. period that each figure was exposed, *S* estimated its relative degree of symmetry and recorded each judgment upon a separate rating scale. If the analytic method was being employed, *S* first indicated the quadrant of poorest symmetry by checking a special diagram at the top of each scale. This technique was designed to reinforce the analytic frame of reference throughout the rating process.

RESULTS

Total distributions.—The mean of each *S*'s distribution of 46 ratings was obtained, and the group means and *SD*'s, based on these individual means, for each experimental group are listed in Table 1.

The statistical significance of the various mean differences was tested by the triple-classification method of the analysis of variance. Bartlett's test for homogeneity of variance produced a chi square of 4.399 with a probability between .80 and .70 that a common population was involved. The variance associated with set, with an F of 26.08 and a p of .001 ($df = 1$), was the only one that proved to be significant. In general, then, it may be said that the influence of the suggestion predominated throughout the experiment, and that it was instrumental in producing judgments of better than average symmetry.

Half-distributions.—In order to investigate the possibility that the anchoring point and method of judgment might have been significant during the early or late stages only of the figure presentations, individual and group means were computed for each half of the distributions (see Table 1).[3]

The triple-classification method of the analysis of variance was again

[3] In order to make these means comparable to those for the total distributions, the individual means were computed by (*a*) obtaining the difference between the sums of the actual and the judged figure values, (*b*) dividing by 23 to get the average difference, and (*c*) adding this average difference to 4.0.

TABLE 1

GROUP MEANS BASED ON THE INDIVIDUAL MEANS OF THE TOTAL DISTRIBUTIONS AND THE TWO HALF-DISTRIBUTIONS*

Group	Variables	Distribution					
		1st Half		2nd Half		Total	
		Mean	*SD*	Mean	*SD*	Mean	*SD*
I	Set, Anchoring Point, and Total Impression	3.76	.65	3.96	.28	3.86	.43
II	Set, Anchoring Point, and Analysis	4.14	.43	4.20	.69	4.17	.53
III	No Set, Anchoring Point, and Total Impression	4.67	.30	4.59	.63	4.63	.41
IV	No Set, Anchoring Point, and Analysis	4.47	.46	4.53	.61	4.50	.44
V	Set, No Anchoring Point, and Total Impression	4.03	.61	3.74	.47	3.89	.43
VI	Set, No Anchoring Point, and Analysis	3.97	.89	3.97	.55	3.97	.66
VII	No Set, No Anchoring Point, and Total Impression	4.54	.51	4.23	.55	4.38	.44
VIII	No Set, No Anchoring Point, and Analysis	4.45	.48	4.47	.52	4.46	.44

* The individual means of the total distributions and two half-distributions may be obtained from Appendices 3 & 4, pages 105–112, of the original dissertation which is on file with University Microfilms, Inc., of Ann Arbor, Michigan.

utilized with each half-distribution processed separately. The set, with F's of 20.83 and 17.12 and p's of .001 ($df = 1$), proved to be highly significant in both halves. The anchoring point, although far from significance in the first half, almost reached the 5% level in the second half-distribution ($F = 3.45$, $df = 1$). The critical interhalf differences, as far as the anchoring point is concerned, are those for Groups I and II where set is also involved. The mean difference of .130 for these two groups produced a t of 2.17 ($.05 > p > .02$). In view of the congruent results obtained in the intrahalf analysis, the increasing influence of the anchoring point may be accepted at the 5% level of confidence.

DISCUSSION

In this exploratory study the distorting influence of the suggestion-induced set was much greater than the stabilizing influence of the anchoring point. Previously cited studies (**1, 2, 3, 7**) have shown that such a suggestion can have an appreciable effect upon perception and judgment. However, they have also found that ambiguity or lack of meaning was a necessary condition. In the present experiment, it was believed that the anchoring point would provide the meaning required to make the situation unambiguous. With it present, *S*s had a figure similar to those being judged, the degree of symmetry of which was known, and continuously available for purposes of comparison.

Furthermore, the stabilizing influence of the anchoring point, as demonstrated in previously cited studies (**4, 6, 8**) should have counteracted the suggestion. Perhaps more or differently placed anchors are necessary when the judges or perceivers have been subjected to false instructions. Johnson has supplied evidence pointing to the fact that the midpoint anchor is not the most effective (**5**). The possibility that the average figure as an anchoring point was forgotten, ignored, or deliberately disregarded does not seem plausible. Many *S*s believed that they disregarded the suggestion, while none of them professed to have done so with the anchor.

The analytic method of judgment was introduced in an attempt to improve upon the preciseness of such subjective measuring when the objects to be judged are complex. It would seem that a more accurate measurement could be made if each component part were referred separately to the continuum rather than the

complex as a whole. The hypothesis to this effect, however, was not confirmed by the results.

There are several factors, other than equivalence of efficiency, that might have contributed to the lack of discrimination between the two methods of judgment. First, there were frequent remarks from total-impression *S*s to the effect that they actually resorted to some degree of analysis. Converse reports from analytic *S*s were obtained much less frequently. These two conditions, therefore, were not mutually isolated. Second, the 5-sec. interval was probably too short for a complete analytic judgment, and some *S*s undoubtedly were somewhat confused.

Summary

Two heretofore separate lines of investigation into the dimensions of perceptual judgment were integrated by examinating the interacting influences of a visual anchoring point and a conflicting mental set upon the formation of a scale of perceptual judgment. The influence of the suggestion-induced set, with or without the presence of the anchoring point, was significant beyond the .1% level of confidence. A gradual increase in the influence of the anchoring point throughout the series of judgments could be accepted at the 5% level of confidence. It was concluded that: (*a*) the influence of a mental set may be a strong determining factor in the perception and subsequent judgment of an unstructured situation; and (*b*) the presence of an anchoring point may be of help in stabilizing the perceptions and judgments, but its influence should not be taken for granted.

References

1. Berenda, R. W. *The influence of the group on the judgments of children.* New York: King's Crown Press, 1950.
2. Farnsworth, P. R., & Beaumont, H. Suggestion in pictures. *J. gen. Psychol.*, 1929, **2**, 362–363.
3. Haggard, E. A., & Rose, G. J. Some effects of mental set and active participation in the conditioning of the autokinetic phenomenon. *J. exp. Psychol.*, 1944, **34**, 45–59.
4. Heintz, R. K. The effect of remote anchoring points upon the judgment of lifted weights. *J. exp. Psychol.*, 1950, **40**, 584–591.
5. Johnson, D. M. A systematic treatment of judgment. *Psychol. Bull.*, 1945, **42**, 193–224.
6. Rogers, S. The anchoring of absolute judgments. *Arch. Psychol., N. Y.*, 1941, No. 261.
7. Sherif, M. An experimental study of stereotypes. *J. abnorm. soc. Psychol.*, 1935, **29**, 371–375.
8. Volkmann, J. The anchoring of absolute scales. *Psychol. Bull.*, 1936, **33**, 742–743.

(Received February 12, 1953)

Journal of Experimental Psychology
Vol. 46, No. 5, 1953

THE EFFECTS OF VARIOUS KINDS OF RELEVANT VERBAL PRETRAINING ON SUBSEQUENT MOTOR PERFORMANCE[1]

DOROTHY ELSEY McALLISTER

State University of Iowa[2]

The purpose of the present investigation was to test several hypotheses relating to the effectiveness of verbal paired-associates pretraining on the learning of a motor task. Recent studies pertinent to the problem have been of two types. In one type, designated as "stimulus predifferentiation," *Ss* have learned during pretraining to associate words or other symbols with stimuli later used in the motor task. In the other type, referred to as "verbal pretraining," *Ss* have learned to associate words or phrases with stimuli which substituted for those later used in the motor task, the words or phrases indicating the correct motor response to be made to each stimulus.

Two investigators (**4, 6**) have reported facilitation of motor performance in a group given stimulus-predifferentiation training as compared with a control group given either no pretraining (**6**) or pretraining with stimuli irrelevant to the motor task (**4**). In another study (**5**), no significant differences were found between a stimulus-predifferentiation group and a group trained with stimuli irrelevant to the motor task. However, both groups performed the motor task significantly better than did a group given no pretraining.

Verbal pretraining has been found in one instance (**1**) to facilitate motor performance as compared to that of a control group given no pretraining, while in another case (**12**) such pretraining had no effect on motor performance in comparison with a control group pretrained on a task unrelated to the motor task.

Several levels of experimental treatment are possible in studies of these types. In an effort to make them explicit and thereby to indicate the appropriate control groups to be utilized, a new terminology is suggested. Three levels of pretraining may be distinguished on the basis of the relevance of the stimuli and response words (or other symbols) to the stimuli and responses of the motor task. They are: (*a*) *Irrelevant:* pretraining involving irrelevant stimuli and irrelevant response words; (*b*) *Relevant S:* pretraining involving relevant stimuli and irrelevant response words; and (*c*) *Relevant S–R:* pretraining involving relevant stimuli and relevant response words. Relevant stimuli are those which are the same as, or reasonable substitutes for, the motor-task stimuli while relevant response words are those which indicate, for the motor task, the appropriate responses to particular stimuli. Each level may be considered as a control for the succeeding level provided that other vari-

[1] This paper is based on a dissertation submitted to the Graduate College of the State University of Iowa in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The work was done under the supervision of Dr. Don Lewis and was supported in part by the U. S. Air Force under Contract No. AF 33 (038)-13214 monitored by the Perceptual and Motor Skills Research Laboratory, Human Resources Research Center. Permission is granted for reproduction, translation, publication, use and disposal in whole or in part by or for the United States Government.

[2] Now at Syracuse University.

ables are adequately controlled. Relevant-S pretraining corresponds to what has previously been called stimulus predifferentiation and relevant-S–R, to verbal pretraining.

The underlying mechanisms accounting for the results of the studies being considered have been indicated elsewhere (**1, 2, 4, 5, 6**), but certain extensions and clarifications are necessary in order to predict the results of the present study. Superior motor performance following relevant-S pretraining as compared with irrelevant pretraining is assumed to result from a transfer of the verbal responses learned to the stimuli of the pretraining task to the stimuli of the motor task. The addition of response-produced cues resulting from such transfer would decrease the similarity among the stimuli of the motor task, and, hence, would decrease the possibility of stimulus generalization over that possible in the control group. This expectation is contingent on the use of a high level of learning for the verbal pretraining. With a low level, the stimuli might elicit incorrect verbal responses with a resultant increase in the similarity of the stimulus complexes and an increase in the amount of generalization. In either case, since some response strength must exist between a stimulus and a response before stimulus generalization can occur, the difference in performance would not be reflected until some amount of motor practice had been given.

Since the transfer of verbal responses presumably results from both relevant-S–R and relevant-S pretraining, any superiority of the former treatment must depend on some additional factors. The explanation involves the assumption that the verbal responses serve as cues, which, as a result of past learning (plus instructions), elicit responses, presumably implicit, which furnish information concerning the correct motor responses. On the assumption that the amount of past learning would vary for different kinds of verbal responses, one would expect such variation to lead to differential amounts of facilitation. If, in addition, in the past history of *S*, response strength had been acquired between the verbal cues and motor responses similar to those required by the motor task, greater facilitation would be expected.

Hypotheses to be Tested

In order to obtain the kinds of pretraining situations required to test the hypotheses, it was necessary to utilize a motor task (apparatus) possessing certain features. The Star Discrimeter, an apparatus patterned after one described by Duncan (**3**), was used. When performing on this device, *S* learns to associate each of six colored light stimuli with one of six horizontal channels, and to move a vertical wobble-stick into the appropriate channel. The six channels, which radiate out from a center opening through which the wobble-stick projects in its centered position, are located at 60° intervals. It is possible to conceptualize the arrangement of the channels in various ways. For example, using a clock analogue, one can think of the channels as pointing toward the even-numbered hours. Other feasible analogues are angular distances in degrees from some point of origin and directions of radiation of the channels from the center opening. Additional features possessed by the discrimeter, which were necessary for the study, are that (*a*) the S–R relationships can be rotated by different amounts, and (*b*) the stimuli can be duplicated on paper for paired-associates learning.

The hypotheses to be tested are stated below. All the predictions were made with respect to performance on the early trials of the motor task. Hypotheses 1, 2, and 4 were suggested by the theoretical considerations stated above, while Hypothesis 3 was suggested by the results of a previous study (**8**). In the absence of

quantitative information about the familiarity of the analogues, personal judgment was used.

Hypothesis 1.—Relevant verbal pretraining of the S–R type serves to facilitate subsequent motor performance over that of a relevant-S control group. Two necessary conditions are: (*a*) that *S*s in the relevant-S–R group be informed concerning the analogue on which the verbal responses of pretraining were based and, (*b*) that *S*s have had previous experience with the analogue.

Hypothesis 2.—The amount of facilitation of motor performance arising from relevant-S–R pretraining depends on the kind of response words (that is, the analogue) utilized in pretraining. A *directions*, a *clock*, and a *degrees* analogue are to be compared. The prediction is that pretraining based on the directions analogue will be most facilitative of motor performance, followed in turn by the clock and the degrees analogues.

Hypothesis 3.—The amount of facilitation arising from relevant-S–R pretraining is inversely related to the amount of clockwise rotation of the S–R relationships from the pretraining to the motor task.

Hypothesis 4.—Relevant-S pretraining serves to facilitate subsequent motor performance over that of an irrelevant pretraining control group. For the reasons stated previously, a difference in performance between the groups is not expected to appear until after a trial or two.

Method

Apparatus.—The adjusted verbal learning apparatus, which has been described in detail elsewhere (11), was used for the paired-associates pretraining. This apparatus allows for the presentation of the paired associates in any desired sequence. Each stimulus appears alone for 2 sec. followed by a 2-sec. presentation of both the stimulus and its associated response.

The Star Discrimeter was used for the motor learning. The response unit of the apparatus, containing the channels and the wobble-stick described above, is placed on a platform adjustable to *S*'s height. A stimulus box is also mounted on the platform at about eye-level and about 30 in. in front of *S*. A circular opening, 2½ in. in diameter and covered with opal glass, is cut in the center of the stimulus box. Six colored-light stimuli are projected one at a time in mixed order through this opening. When the wobble-stick is moved all the way into a correct channel, a 50-point stepping switch operates to present another colored light. A variety of interconnections between the lights and channels are possible.

Two microswitches, placed ½ in. and 2½ in. from the entrance to each 3½-in. channel and another placed at the end of each, permit a count of the number of shallow and deep errors and correct responses, respectively. The deep errors and correct responses are recorded on counters while the number of shallow errors is obtained through an electrical polygraph.

Subjects.—The *S*s were 98 male students from the elementary psychology class at the State University of Iowa. All *S*s were able to read correctly the numbers on Ishihara color cards 1–9 and 12–13.

Design and procedure.—The *S*s were assigned at random to seven groups of 14 each. The first part of the experimental session was devoted to the paired-associates learning of a list of six stimulus-response pairs by the adjusted learning method. The initial criterion of learning was five correct anticipations of each response word within a time limit of 14 min. When this criterion was met, *S* remained seated and read magazines until 15 min. had elapsed from the time the learning was begun. Then the entire list of paired associates was presented to a criterion of one correct anticipation of each pair.

The type of pretraining given to the various groups is indicated in the bottom row of Table 1. Groups 3, 4*a*, 4*b*, 4*c*, and 5 were all given relevant-S–R pretraining. The stimuli were circular color patches (19 mm. in diameter), made with water colors, which matched as closely as possible the six colors on the discrimeter. The names of the colors are shown in the first column of Table 1 and the response words to be associated with them are shown in Columns 3, 4, and 5. The response words for Group 3 were based on a degrees analogue of the channels of the discrimeter (with zero degrees at the farthest channel and with a clockwise progression); for Groups 4*a*, 4*b*, and 4*c*, on a clock analogue; for Group 5, on a directions analogue. The stimuli for Group 2 were color patches, while the stimuli for Group 1 were words which were irrelevant to the motor task. The same list of response words, also irrelevant to the motor task, was used for Groups 1 and 2. These lists are presented in the second and last two columns of Table 1.

Immediately after completion of the criterial trial and approximately 16 min. after the beginning of the verbal learning, *S* was taken to another building for practice on the discrimeter. Each *S*

TABLE 1

STIMULI AND ASSOCIATED RESPONSE WORDS USED DURING VERBAL PRETRAINING AND TYPE OF PRETRAINING FOR EACH GROUP

Color of Stimulus Patch	Response Words				Group 1	
	Group 2	Group 3	Groups 4*a*, 4*b*, and 4*c*	Group 5	Stimulus	Response
Blue-green	Valiant	0 Degrees	12 O'clock	Forward	Inert	Valiant
Orange	Haggard	60 Degrees	2 O'clock	Right-forward	Pious	Haggard
Purple	Uncouth	120 Degrees	4 O'clock	Right-backward	Rustic	Uncouth
White	Sullen	180 Degrees	6 O'clock	Backward	Jumbo	Sullen
Yellow	Crafty	240 Degrees	8 O'clock	Left-backward	Earthly	Crafty
Red	Fitful	300 Degrees	10 O'clock	Left-forward	Duplex	Fitful
Type of pretraining	Relevant S	Relevant S–R	Relevant S–R	Relevant S–R	Irrelevant	

was given 45 20-sec. trials, with 10-sec. intertrial intervals and a 1-min. rest period after each block of 15 trials.

Groups 1, 2, 3, 4*a*, and 5 practiced on Task A. On this task the stimulus-response relationships learned by Groups 3, 4*a*, and 5 during the pretraining were entirely appropriate provided the channels were conceptualized in terms of the proper analogue. Group 4*b* practiced on Task B, and Group 4*c*, on Task C. Tasks B and C involved stimulus-channel relationships which were rotated 4 hr. and 8 hr., respectively, in a clockwise direction, from those obtaining on Task A.

All the relevant-S–R groups were instructed to conceptualize the placement of the channels in terms of the appropriate analogue. The rotation groups were instructed regarding the amount of clockwise rotation. The relevant-S–R and relevant-S groups were told that the colors on the discrimeter were the same as those used during the pretraining. The remainder of the instructions were routine and were identical to those given to the irrelevant pretraining group.

RESULTS

Verbal Pretraining

The total number of presentations to the pretraining criterion of six correct anticipations of each of the paired associates was obtained for each *S*. Using a 5% coefficient of risk, which was adopted for all statistical tests, a Bartlett test indicated that the variances of these measures were nonhomogeneous among the groups ($\chi^2 = 14.165$ with 6 *df*). A nonparametric test described by Mood (**10**, p. 398) was used to test the hypothesis that the measures were drawn from the same population. A χ^2 value of 8.87 with 4 *df* was obtained when Groups 4*a*, 4*b*, and 4*c* were combined; a value of 9.07 with 6 *df*, when these groups were considered separately. The hypothesis was tenable in both cases. Therefore, it was assumed that any differences observed in motor performance were attributable only to the variations in the experimental treatments.

Motor learning

The raw data consisted of number of correct responses and number of errors per trial, from which group means were obtained and plotted. Although the deep error curves tended to reach a somewhat lower level than the shallow error curves as practice continued, there seemed to be no differences in their trends, so the analysis was limited to total errors.

Simple analysis of variance was used to evaluate the means of correct responses and of errors

over Trials 1–3 (or 2–4 for Hypothesis 4) unless otherwise indicated. The within-groups variance estimate was used in computing the error term for the *t* tests. The correct response data over the 45 trials and the error data over the first 15 trials were analyzed in the manner indicated by Lindquist (7) for the between-*S* component of the total sum of squares in a Type I design. Differences between pairs of general means were evaluated with *t* tests, the between-*S*s-within-groups variance estimate being used to compute the error term.

The *t* tests for the initial trials were one-tailed since predictions had been made, while those involving general means were two-tailed.

The assumption of homogeneity of variance was tested with a Bartlett test prior to each application of analysis of variance. In the absence of adequate statistical tests, the assumptions of random sampling and normality were judged tenable on the basis of the sampling procedure and observation of the data, respectively.

In the following paragraphs, the experimental findings and their statistical analyses are presented separately for each hypothesis.

Hypothesis 1.—As may be seen by an inspection of the means and the *t*'s shown in Table 2, Group 2 (relevant S) made, with one exception, significantly fewer correct responses over the 45 trials and significantly more errors both at the outset of practice and over a number of trials than did either Group 3, 4*a*, or 5 (relevant S–R).

Since the variance estimates based on the means of correct responses over the first three trials were found to be nonhomogeneous ($\chi^2 = 10.397$ with 3 *df*), the Mann-Whitney *U* test (9) was used to evaluate the differences among these measures. The resulting *z* values led to a rejection of the null hypothesis at better than the .1% level of confidence for Groups 2 and 4*a* and also for 2 and 5 ($z = 3.42$ and 3.61, respectively). Contrary to prediction, the performance of Group 3 was not superior to that of Group 2.

It is apparent from Fig. 1, in which means of correct responses are plotted for each of the 45 trials, that Group 2 performed less proficiently than did Groups 4*a* and 5 and quite comparably to Group 3. The striking differences between the error curves, plotted in Fig. 2, which tend to persist throughout most of the practice, indicate that relevant-S–R pretraining led to superior motor performance in all cases. With both measures the several curves tend to converge toward the end of practice.

Hypothesis 2.—The means of correct responses on the first three trials and the general means for Groups 3, 4*a*, and 5, which may be seen in Table 2, differed significantly among the three groups ($F = 9.79$ and 4.93 with 2 and 39 *df* for initial trial means and

TABLE 2

MEANS OF CORRECT RESPONSES AND ERRORS FOR SELECTED TRIALS OF GROUPS 2, 3, 4*a*, AND 5, WITH RELEVANT *F* AND *t* VALUES

Measure	Means for Groups				*F*	*t* Ratios for Groups		
	2	3	4*a*	5		2 vs. 3	2 vs. 4*a*	2 vs. 5
Correct responses								
Trials 1–3	4.24	4.07	7.81	9.07				
Trials 1–45	13.36	13.88	15.55	16.45	6.06**	.63	2.64*	3.72**
Errors								
Trials 1–3	19.93	5.62	5.26	6.98	62.73**	11.45**	11.74**	10.36**
Trials 1–15	17.28	4.30	3.98	5.27	34.36**	8.37**	8.58**	7.75**

* $.05 > p > .01$. For each *F*, $df = 3$ and 52; for each *t*, $df = 52$.
** $p < .01$.

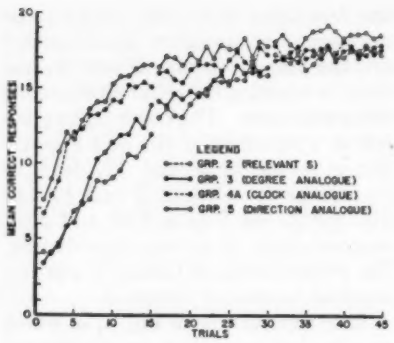


FIG. 1. Means of correct responses for each of the 45 trials on the Star Discrimeter for Groups 2, 3, 4*a*, and 5

general means, respectively). The difference in both the initial trial and the general means of the groups with the degrees and the directions analogues were significant ($t = 4.274$ and 3.096, respectively). The difference in initial trial means of the groups with the degrees and the clock analogues was significant ($t = 3.197$) while the difference in general means approached significance ($t = 2.012$). No significant differences in means were found between the groups having the clock and the directions analogues ($t = 1.077$ and 1.084 for initial trial and general means, respectively). The trends of the curves fell in the predicted order, as may be seen in Fig. 1.

The hypothesis that mean errors on the first three trials would be lowest for Group 5, next for Group 4*a*, and highest for Group 3 was not supported by the data. As may be seen in Table 2, the mean for Group 5 was highest while those for Groups 3 and 4*a* were about the same. The general means over the first 15 trials did not differ significantly, the F ratio being less than one. As shown in Fig. 2, the trends of the curves were similar throughout practice.

Hypothesis 3.—While performance of the two rotation groups was found to be quite comparable throughout practice, Group 4*b* tended to make more errors over the first 20 trials. However, the differences were too inconsequential to warrant presentation of the data. Both rotation groups made about the same number of correct responses but decidedly fewer errors than did the relevant-S group, indicating that even with the rotation, relevant-S–R pretraining led to more efficient performance. Group 4*a* made significantly more correct responses at the outset of practice than either Group 4*a* or 4*b* ($F = 8.53$ with 2 and 39 *df*; $t = 3.899$ and 3.172, respectively) and fewer errors but the latter differences lacked significance.

Hypothesis 4.—The means of correct responses over Trials 2–4 were practically identical for Groups 1 and 2; 4.90 and 4.86, respectively. The curves for correct responses in Fig. 3 show that the two groups performed similarly throughout practice, although the curve for Group 2 fell quite consistently below that for Group 1. The difference between the general means over the 45 trials was not significant ($F = 1.10$ with 1 and 26 *df*).

Contrary to expectations, the error measure indicated that the performance of Group 2 was inferior to that of

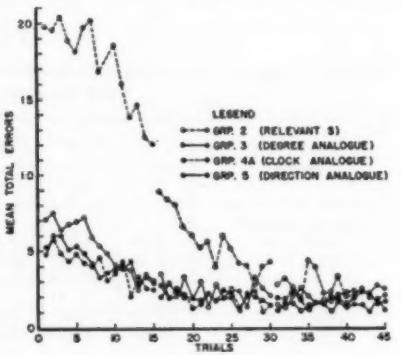


FIG. 2. Means of errors for each of the 45 trials on the Star Discrimeter for Groups 2, 3, 4*a*, and 5

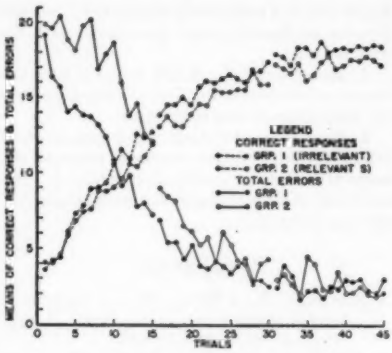


FIG. 3. Means of correct responses and of errors for each of the 45 trials on the Star Discrimeter for Groups 1 and 2

Group 1 (see Fig. 3). The means over Trials 2–4 were 15.38 for Group 1 and 19.67 for Group 2. A *t* test yielded a ratio of 2.665 with 26 *df*, which allowed for a rejection of the null hypothesis beyond the 2% level of confidence. The general mean of errors over the first 15 trials for Group 2 (17.29) was significantly greater than that for Group 1 (12.37). The obtained value of *F* (4.49 with 1 and 26 *df*) fell beyond the 5% level of confidence.

DISCUSSION

Support for Hypotheses 1 and 2 was furnished by the present results. Relevant-S–R pretraining facilitated later motor performance (as indicated by a greater number of correct responses and/or fewer errors) as compared to relevant-S pretraining. The amount of facilitation, in terms of the correct response measure, depended on the analogue upon which the pretraining response words were based; that is, the performance of the groups trained with the directions and the clock analogues was superior to that of the group trained with the degrees analogue. However, the predicted superiority arising from the use of the directions analogue as compared to the clock analogue was not upheld. It had been assumed that the directions, but not the clock, response words had been associated in the past history of *S* with movements similar to those required on the discrimeter, and thus some response strength between the transferred verbal cues and the motor responses would be expected at the outset of practice. In all other respects the use of the two analogues was deemed to be equally effective. The clock analogue was expected to be superior to the degrees analogue as a result of a presumed greater familiarity due to past experience. With few exceptions *S*s using the degrees analogue reported having had little previous experience with degrees. In light of the inferior performance of this group, the rationale seems warranted.

Hypothesis 3 was rejected since, contrary to expectations, performance did not vary with the amount of rotation of the S–R relationships from the pretraining to the motor task. McFann (8) found, in a motor learning study on the Star Discrimeter, that interpolated learning involving a 120° rotation of the stimulus-channel relationships from those obtaining in original learning resulted in poorer performance than did a 60° rotation. The difference in results may be due to differences in the type or amount of prior training or possibly to the fact that different amounts of rotation were involved in the two studies.

No support was obtained for Hypothesis 4. Although the number of correct responses was about the same, a greater number of errors occurred following relevant-S pretraining than following irrelevant pretraining. This finding is contrary to the results reported by Dysinger (4), but the results in terms of the correct response measure are in agreement with those reported by Farber and Murfin (5).

There are at least two possible explanations of the present results. First, the verbal-learning criterion may have been too low, so that for the relevant-S group the stimuli of the motor task elicited incorrect verbal responses with a resultant increase in generalization and a consequent increase in errors. A second explanation is in terms of the conditions of pretraining. The 2-sec. rate of presentation used during the verbal learning required the *S*s to respond rapidly to the

stimuli. If a set to respond rapidly to colors were established in the *Ss* of the relevant-S group and if this set transferred to the motor task, these *Ss* would be expected to respond rapidly during the motor practice. The *Ss* in the irrelevant pretraining group, given no such experience with colors, would not be expected to acquire such a set. Since the majority of the additional responses made by *Ss* in the relevant-S group during the early stages of practice (before the S–R relations were learned) would be incorrect, differences in number of correct responses might not be expected. An experimental test of this hypothesis could be made by comparing the motor performances of several groups given relevant-S pretraining with varying rates of presentation.

Summary

Seven groups of 14 *Ss* were given one of three kinds of paired-associates pretraining prior to learning to associate each of six colored-light stimuli with one of six horizontal channels and to move a vertical response rod into the appropriate channel (Star Discrimeter). The six channels, which radiate out from a center opening, are located at 60° intervals. Measures of correct responses and errors were obtained on each of 45 20-sec. trials.

One group of *Ss*, during *irrelevant* pretraining, learned to associate pairs of words which were unrelated to the motor task. During *relevant-S* pretraining, another group learned to associate words unrelated to the motor task with stimuli which simulated those used in the motor task. Five separate groups, during *relevant-S–R* pretraining, learned to associate the substitute stimuli with response words based on one of three analogues of the channels on the discrimeter: a "degrees," a "clock," or a "directions" analogue. Two groups, given pretraining with the clock analogue, practiced on a motor task involving either a 4-hr. or an 8-hr. rotation of the S–R relationships from those obtaining during pretraining. The relevant-S–R groups were instructed as to how the pretraining was related to the motor task.

The principal findings were as follows:

1. Relevant-S–R pretraining facilitated subsequent motor performance.

2. The use of either the directions or the clock analogue led to better performance in terms of the correct response measure than did the use of the degrees analogue. The difference between the groups with the directions and the clock analogues was not statistically significant. No differences among the groups were found with the error measure.

3. Motor performance did not vary with the amount of rotation of the S–R relationships from the pretraining to the motor task.

4. Pretraining of the relevant-S type resulted in a significantly greater number of errors on the motor task than did irrelevant pretraining. No significant difference was found with the correct response measure.

References

1. Baker, K. E., & Wylie, R. C. Transfer of verbal training to a motor task. *J. exp. Psychol.*, 1950, **40**, 632–638.
2. Dollard, J., & Miller, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
3. Duncan, C. P. Transfer in motor learning as a function of degree of first-task learning and inter-task similarity. *J. exp. Psychol.*, 1953, **45**, 1–11.
4. Dysinger, D. W. An investigation of stimulus pre-differentiation in a choice discrimination problem. Unpublished doctor's dissertation, State Univer. of Iowa, 1951.
5. Farber, I. E., & Murfin, F. L. Performance set as a factor in transfer of training. Paper read at Midwest. Psychol. Ass., Chicago, April, 1951.
6. Gagné, R. M., & Baker, K. E. Stimulus pre-differentiation as a factor in transfer of training. *J. exp. Psychol.*, 1950, **40**, 439–451.
7. Lindquist, E. F. *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin, 1953.
8. McFann, H. H. Effects of response alteration and different instructions on proactive and retroactive facilitation and interference. *J. exp. Psychol.*, 1953, **46**, in press.
9. Mann, H. B., & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Statist.*, 1947, **18**, 50–60.
10. Mood, A. M. *Introduction to the theory of statistics*. New York: McGraw-Hill, 1950.
11. Shephard, A. H., & Lewis, D. A new device for presenting paired-associates in verbal-learning. *Amer. J. Psychol.*, 1950, **63**, 600–606.
12. Spieth, W. A study of the effects of verbal pre-training on the subsequent learning of two mutually antagonistic motor tasks in alternating practice. Unpublished master's thesis, State Univer. of Iowa, 1951.

(Received February 16, 1953)

Journal of Experimental Psychology
Vol. 46, No. 5, 1953

BILATERAL REMINISCENCE IN PURSUIT-ROTOR LEARNING AS A FUNCTION OF AMOUNT OF FIRST-HAND PRACTICE AND LENGTH OF REST[1]

MARTY R. ROCKWAY[2]

Northwestern University

Reminiscence has been defined as "an improvement in performance, as shown by some measure of ability to recall at some time after the original practice, without (any) intervening practice" (7, p. 314). In conventional studies of reminiscence employing pursuit-rotor learning, *Ss* use the same hand for both pre- and postrest practice, and, commonly, the performance of a no-rest control group is the base line from which the improvement over rest is calculated.

Recently a phenomenon similar to conventional reminiscence in pursuit-rotor learning has been reported by Ammons and Ammons (4), Kimble (19), and Irion and Gustafson (16). In both the Kimble and the Irion and Gustafson experiments two groups of *Ss* were given 5 min. of massed practice on the pursuit rotor. Following this, *Ss* in one group immediately received an additional 5 min. of practice with the other hand, while *Ss* in the second group rested for 5 min. before changing hands. The experimental data revealed that during other-hand practice the performance of the group that rested was significantly superior to that of the no-rest group.

These demonstrations of bilateral "reminiscence" may prove to be of considerable importance in providing a technique for studying the nature and locus of distribution of practice and reminiscence effects. Previous studies of these phenomena have justified the postulation of a response-produced inhibitory factor which temporarily depresses performance during motor learning. This factor, which has been identified with Hull's (14) concept of reactive inhibition (I_R), is assumed to accumulate under conditions of continuous practice and to dissipate with rest. Hull assumes that the locus of the performance decrement resulting from the accumulation of I_R is confined to the peripheral effectors involved in responding. However, the evidence of bilateral reminiscence suggests that inhibitory effects may be more general; i.e., the inhibition generated by one response system may depress the performance of another, apparently different, response system.

Despite the apparent similarity between conventional and bilateral reminiscence, the evidence is not sufficient to conclude that they both can be adequately explained by a single theoretical construct. A first step in the direction of establishing such an explanatory identity would be to show that the two effects are similar functions of the same variables.

[1] This report is a portion of a dissertation submitted to the Department of Psychology of Northwestern University in partial fulfillment of the requirements for the Ph.D. degree. The author wishes to express his appreciation to Dr. Carl P. Duncan for his assistance in the conduct of this research.

[2] Now at the Psychology Branch, WCRDP-5, Aero Medical Laboratory, Directorate of Research, Wright-Patterson AFB, Ohio.

Ammons (2), Bell (5), Kimble and Horenstein (20), Melton (21), and Irion (15) have published pursuit-rotor studies concerned with the relationship between the amount of conventional reminiscence and the length of the interpolated rest

interval. In general, the results show that reminiscence is a negatively accelerated increasing function of the length of the rest interval, reaching a maximum after about 5-min. rest. Both Ammons and Irion also investigated reminiscence as a function of the duration of prerest practice. Their findings indicated that reminiscence at first increases and then decreases with increasing amounts of prerest practice. The greatest gain over rest occurred after 8 to 10 min. of practice.

A comparison of the effects of varying amounts of interpolated rest on conventional and bilateral reminiscence has very recently been reported by Grice and Reynolds (12). This investigation consisted of two parallel experiments in which rest periods of 10 sec., 30 sec., and 1, 3, 5, and 10 min. were introduced after 15 30-sec. trials on the pursuit rotor. In one experiment *Ss* used the left hand before and after rest, while in the other experiment they used the right hand before rest and left hand afterwards. Reminiscence was an increasing function of amount of rest for both conditions. Although the gains over rest of the left-left groups were larger than those of the right-left groups, the hypothesis was tenable that the functions were the same except for asymptotic value.

In addition to its theoretical significance, the phenomenon of bilateral reminiscence points to the need for controlling the time intervals in bilateral transfer-of-training research. For example, Cook (9) states that bilateral transfer is an increasing function of the amount of initial practice. However, demonstrations of the transfer of inhibition suggest that the shape of this function is dependent upon both the intertrial rest interval during initial practice and the rest interval separating initial and final practice. So, it would appear that any precise description of the relationship between bilateral transfer and initial practice is incomplete unless the rest intervals are also specified.

The purpose of the present study is to determine the amount of bilateral reminiscence in pursuit-rotor learning as a joint function of the duration of first-hand (FH) practice and the length of the rest interval between FH and second-hand (SH) practice. It is convenient to consider this problem within the framework of bilateral transfer of training. Hence, if we accept Cook's generalization regarding the relationship between bilateral transfer and initial practice, and if we anticipate our results by assuming at least a gross similarity in the functions of conventional and bilateral reminiscence, then we are led to the following prediction: The level of SH performance is an increasing function of the amount of FH practice, and the rate of rise of this function becomes greater with increasing lengths of interpolated rest.

Method

Apparatus.—The apparatus was a modification of the Koerth pursuit rotor. The turntable was 28.6 cm. in diameter and rotated clockwise at a speed of 60 rpm. A circular brass target 1.74 cm. in diameter was set flush with the surface of the turntable, with the center of the target 7.8 cm. from the center of the turntable. During practice *S* attempted to keep a hinged brass stylus on the rotating target. Contact between stylus and target activated one of two .01-sec. Standard Electric stop clocks. Scores were recorded after each 30 sec. of practice by manually switching from one clock to the other every 30 sec., so that while *S* was cumulating a score on one clock the other was being read and reset. Length of trials and rest intervals were timed by means of two Kodak spring-wound timers.

Subjects.—Buxton and Grant (8) have demonstrated significant performance differences between men and women in motor learning; therefore, because of greater availability, only women were used in the present experiment. Three hundred *Ss*, 25 per group, were assigned at random to each of the 12 conditions. All *Ss* were enrolled in the introductory social science course or one of the undergraduate psychology courses at Northwestern University. Students in these courses may earn one point of examination credit for each experimental session in which they serve. None of the *Ss* had had previous experience with pursuit-rotor learning.

Thirty *Ss* were eliminated from the experiment on the basis of several criteria, viz.: (*a*) not making contact with the target during the first minute of FH practice (7 *Ss*); (*b*) violating the instructions regarding proper performance of the task for three successive trials during either

FH or SH practice (6 *Ss*); (*c*) answering any of the questions regarding hand preference in a manner suggesting ambidexterity (15 *Ss*); and (*d*) recording difficulties (2 *Ss*). The criterion regarding ambidexterity effected the elimination of many left-handed *Ss*, so that the records of only nine left-handed *Ss* were included in the data.

Procedure.—Without any pretraining trials 12 groups of *Ss* received either 1, 2, 3, or 5 min. of FH practice using their preferred hand, followed by a rest of 0, 2, or 5 min.; then all *Ss* engaged in 5 min. of SH practice using their nonpreferred hand. Practice was continuous, except for interpolated rest, and scores were recorded at the end of each 30-sec. interval.

All *Ss* were first asked a number of questions to determine handedness. Following this, every *S* was read appropriate instructions and given a brief demonstration of proper and improper pursuit-rotor performance by *E*.

During the first minute of FH practice any errors in the method being used by *S* were corrected verbally without interruption of practice. Thereafter, until completion of the experiment no further comments regarding performance adequacy were made by *E*.

At the end of FH practice the rotor was stopped. The *Ss* in the no-rest conditions were told to shift the stylus quickly to the other hand. As soon as the shift was accomplished the rotor was started again. This transfer interrupted the continuity between FH and SH practice for approximately 6 sec. The *Ss* in the rest conditions were instructed to leave the rotor and sit down in a chair facing away from the apparatus. Their rest period was spent in reading magazines or conversing with *E*. Fifteen seconds prior to the end of the rest period they were told: "Return to the rotor, pick up the stylus with your other hand, and be ready to resume practice on signal."

Results

The group means of the basic measures of FH and SH performance (time on target) are presented in Table 1. Since all *Ss* had at least the first two trials of FH practice in common, the total time on target during this period was used as an indicator of initial proficiency (*FH score*). The basic measure of SH performance was the total time on target during the first five trials of postrest practice (*SH score*). This measure was selected because, while it preserves initial postrest performance differences among groups, it is more reliable than earlier scores encompassing fewer trials.

TABLE 1

Mean Total Time on Target in Seconds for Each Condition

Conditions		First-Hand Trials 1–2	Second-Hand Trials 1–5	
FH Practice (Min.)	Rest (Min.)		Obtained	Adjusted*
1	0**	3.74	6.02	5.81
1	2	4.59	8.08	6.75
1	5	3.29	7.98	8.34
2	0**	3.53	5.47	5.52
2	2	3.50	10.17	10.25
2	5	3.97	11.46	10.94
3	0**	3.40	5.55	5.76
3	2	2.83	7.61	8.57
3	5	3.43	11.24	11.42
5	0**	4.56	7.73	6.44
5	2	2.54	10.44	11.78
5	5	3.45	17.28	17.43

* Adjusted by means of the average regression of second-hand Trials 1–5 on first-hand Trials 1–2.
** Actually about 6 sec.

The FH scores were subjected to a simple analysis of variance. This analysis was undertaken to determine whether random assignment of *Ss* had resulted in groups differing in initial proficiency. The obtained F ratio was less than one, indicating that the groups did not differ significantly with respect to this measure.

The variances of the 12 groups of SH scores were tested for homogeneity by Bartlett's test. The obtained chi square of 41.26 indicated that the variances were heterogeneous; with 11 *df*, a chi square of 24.72 is significant at the 1% level. Further examination of the data also revealed a tendency toward proportionality of the group means and *SD*'s, so the SH scores were transformed to a logarithmic scale. In this case, the transformation was to log $Y + 1$; the constant was added to avoid negative values.

Application of Bartlett's test to the transformed SH scores produced a chi square of 18.76; with 11 *df*, a chi square of 19.67 is required for significance at the 5% level.

The conditions of the experiment were chosen so that the resulting post-rest data would be amenable to a two-classification, 3 × 4 factorial analysis of variance. This analysis of the transformed SH scores is summarized in Table 2. Since neither set of values of the independent variables can be considered a random sample, the mean square within groups is the appropriate error term for all tests of significance. Length of rest is significant far beyond the 1% point; duration of FH practice is significant beyond the 1% point; and the interaction of rest and practice is significant between the 1% and 5% points. These *F*'s may be interpreted as indicating that FH practice and rest influence SH performance, but that the effect of manipulating one variable is dependent upon the level of the other variable.

Although the groups did not differ significantly on the measure of initial proficiency, there was considerable variation among the mean FH scores. For instance, an examination of Table 1 reveals that the mean for the 5-min. practice, 2-min. rest group was only 2.54 while that for the 1-min. practice, 2-min. rest group was 4.59. Though not statistically significant, these initial differences may tend to obscure the true effects of the main variables. This confounding will exist to the extent that there is a predictable relationship between the FH and SH scores which is independent of experimental treatments.

TABLE 2

ANALYSIS OF VARIANCE OF TRANSFORMED SCORES (LOG $Y + 1$) OF TOTAL TIME ON TARGET DURING SECOND-HAND TRIALS 1–5

Source of Variation	*df*	Mean Square	*F*
Rest interval	2	1.6397	9.85**
First-hand practice	3	.9833	5.90**
Interaction	6	.3653	2.19*
Within groups	288	.1664	
Total	299		

* $p < .05$.
** $p < .01$.

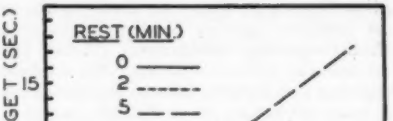


FIG. 1. Adjusted mean total time on target during the first five trials of second-hand practice as a function of amount of first-hand practice. Each curve represents a different length of interpolated rest.

An attempt was made to compensate for the initial differences among groups by adjusting the mean SH scores. The adjustment was made using the procedure outlined by Snedecor (**22**, p. 321 ff.). This technique employs the average regression coefficient of SH scores on FH scores to predict what the SH means would have been if all groups had been equally proficient during the first two trials of FH practice. This regression coefficient, which is an average of the coefficients of the individual groups, is identical with the within-groups regression coefficient computed in an analysis of covariance. The adjusted SH means are presented in Table 1.

An inspection of the adjusted functions in Fig. 1 aids in the interpretation of the significant *F* ratios obtained

by analysis of variance. The general positive trend of the SH means with increasing amounts of FH practice and the upward displacement of the curves with increasing lengths of rest reflect the highly significant F ratios for these variables. The significant F for the interaction of rest and practice is illustrated by the slope differences among the individual curves; i.e., the steepness of the functions in Fig. 1 increases with longer lengths of interpolated rest.

The operational definition of bilateral reminiscence adopted in this study is similar to that used for conventional reminiscence, except, of course, for the changing of hands after rest. The gain of each rest group was measured as the difference between its adjusted mean SH score and the adjusted mean SH score of the no-rest group with a corresponding amount of FH practice. These gains are plotted in Fig. 2. Since only four rather irregular points determined each reminiscence function, it was not deemed worth while to attempt a precise mathematical fit. However, it is obvious that, in general, bilateral reminiscence increases with increasing amounts of FH practice and interpolated rest.

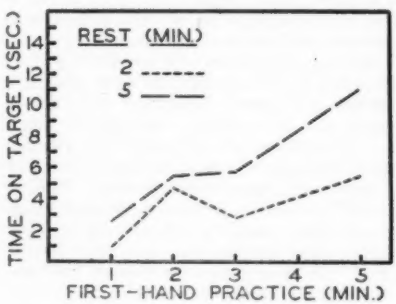


FIG. 2. Reminiscence (time on target) as a function of amount of first-hand practice for 2 and 5 min. of interpolated rest

DISCUSSION

It was predicted that, within the limits of this investigation, SH performance is an increasing function of FH practice, and the rate of rise of this function becomes greater with an increase in the time between FH and SH practice. The results clearly verified the prediction, and, by so doing, they also tend to substantiate the assumptions involved in its formulation. These assumptions, which may be considered as premises from which the experimental hypothesis was deduced, are as follows: (*a*) bilateral transfer is an increasing function of initial practice; (*b*) bilateral reminiscence is an increasing function of initial practice; and (*c*) bilateral reminiscence is an increasing function of the time separating initial and final practice.

Regarding the first assumption, the absolute amount of bilateral transfer of training could not be estimated from the data of this study, since an appropriate control group was not included. However, the subtraction of a constant, equal to the mean of a control group, from each SH mean would not have changed the rank order or the magnitude of the differences among the experimental groups. Therefore, it seems obvious that the conclusions, with respect to the effects of FH practice and rest, are equally valid whether stated in terms of absolute amount of bilateral transfer or level of SH performance.

Actually the primary empirical interest of this study was not in bilateral transfer per se, but, rather, in the determination of the bilateral reminiscence functions. Therefore, despite the fact that the assumptions regarding the shapes of these functions were supported by verification of the experimental hypothesis, a more direct determination was desired. The actual reminiscence measures plotted in Fig. 2, although irregular, seem to indicate that bilateral reminiscence is an increasing function of both FH practice and interpolated rest. Thus, for the range of values examined, bilateral and conventional reminiscence appear to be at least generally similar functions of the

same variables. With regard to the influence of rest, it should be noted that these data corroborate the more extensive findings of Grice and Reynolds (12).

Obviously, a demonstration of functional similarity does not demand that both reminiscence phenomena be subject to the same explanation. On the other hand, it would seem advisable to examine the feasibility of interpreting the two effects in terms of existing concepts before postulating new or different ones. Most currently popular explanations of conventional reminiscence in motor learning (1, 17) embody theoretical constructs essentially similar to Hull's I_R. Therefore, in subsequent paragraphs we shall consider the adequacy of I_R, or, rather, certain characteristics attributed to I_R, in explaining findings of both bilateral and conventional reminiscence.

Effector localization of I_R.—For Hull at least, I_R carries with it the implication of an inhibitory substance localized in the responding effectors (14, p. 281). The evidence of bilateral reminiscence would at first seem to deny this interpretation, until, as Grice and Reynolds observed, the relative complexity of pursuit-rotor performance is remembered. This task involves not only the more obvious shoulder, arm, and hand effectors, but many others as well. Therefore, the gains over rest might well be explained as a recovery from inhibition "localized" throughout the body.

Practically speaking, a localization hypothesis of inhibitory effects is not amenable to a precise experimental test. One reason for this is that, at present, there is no practicable method of completely immobilizing all but a particular response system. This being the case, it can always be argued that other systems participated, albeit subliminally, during the performance of almost any given task. However, if effector I_R is accepted as the *sole* determiner of reminiscence effects, one might expect that the greatest amount of inhibition would accrue in the most active response mechanisms. For example, during pursuit-rotor practice, the arm and hand holding the stylus appear to be most vigorously employed—especially by naive *S*s. Therefore, when hands are switched, a relatively large proportion of total I_R should not transfer. This leads to the prediction that, for pursuit-rotor learning, conventional reminiscence is considerably greater than bilateral reminiscence.

The only published study which permits even a crude test of the above prediction is that of Grice and Reynolds (12). Their data did indicate differences between the two phenomena, in the predicted direction, for comparable conditions of FH practice and rest. Still, these differences were not very great, and perhaps they can be attributed, at least partially, to the relatively poor SH performance of the groups which changed hands. Concerning the relationship between performance level and conventional reminiscence, Buxton says, "In pursuit rotor learning, the absolute amount of reminiscence increases, then decreases, as level of mastery before a constant rest-interval rises" (6, p. 180).

Another bit of evidence which seems to weigh against a purely effector localization theory of temporary inhibition has been supplied by Ammons (3). During 2 min. of prepractice activity, *S*s were blindfolded and they followed with the index finger a small rivet head set in the rotor at the same distance from the center as the regular target on the pursuit rotor. Now, apparently, the movements involved in this activity are quite similar to those made during actual practice. Yet, when compared to a control group, this activity did not significantly inhibit or facilitate pursuit-rotor performance at any time during later practice. These results led Ammons to conclude that temporary work decrement (Hull's I_R) is mainly central rather than peripheral in nature.

It would appear that the findings just reviewed do not support the assumption that the performance decrement observed under conditions of massed practice is confined to the peripheral effectors. Nevertheless, it should be pointed out that arguments against the ancillary hypothesis of effector localization are not really crucial to a reactive inhibition type theory. The operational validity of I_R and its usefulness as a logical construct are quite independent of attendant physiological speculations. However, there is evidence which suggests that some modification of the explanation of even conventional reminiscence in terms of I_R is necessary.

Response effortfulness as a determiner of I_R.—According to Kimble (18), the more *effortful* a response, the greater the amount of I_R that is generated. For example, pursuing a target rotating at 80 rpm should expend more physical energy than pursuing the same target rotating at 50 rpm for an equal period of time. Thus, if Kimble's proposition is true, the 80-rpm group should develop more I_R and, as a consequence, display more reminiscence than the 50-rpm group. An investigation by Helmick (13) provides evidence which is pertinent to this hypothesis. He found that, with the pursuit rotor, neither variations in speed of rotation nor size of target significantly affected the amount of reminiscence.

Two other experiments also furnish data regarding the relationship between the degree of task effortfulness and the amount of reminiscence in motor learning. Eckstrand (10), by manipulating spring tension, varied the force required to depress the stylus in pursuit-rotor learning. He found that no significant differences in reminiscence resulted from these variations. Ellis, Montgomery, and Underwood (11), using a block-turning task, altered effortfulness by varying the work-surface height. They also failed to find significant differences in reminiscence.

The various findings presented and discussed in this report suggest a need for much additional experimental study of reminiscence phenomena. In view of this, perhaps it would be wise to defer highly theoretical speculations in this area. A more fruitful approach might be to direct most of our immediate efforts toward the discovery and systematic investigation of the empirical variables which govern the transfer of reminiscence effects. Preliminary studies, like the present one, could reveal the relevant variables and suggest the general form of the relationships. These explorations could then be followed by attempts at very reliable determinations of the functions over a wide range of values of the independent variables. Such careful and complete descriptions should provide a more reliable factual foundation upon which to construct subsequent theory.

Summary

Twelve groups of 25 *Ss* received either 1, 2, 3, or 5 min. of first-hand (FH) practice on the pursuit rotor with their preferred hand; this was followed by a rest of 0, 2, or 5 min.; then all *Ss* engaged in 5 min. of second-hand (SH) practice using their nonpreferred hand.

Performance during SH practice was an increasing function of amount of FH practice and length of rest between FH and SH practice. The slope of the FH practice-SH performance function increased with increasing lengths of interpolated rest.

Bilateral reminiscence, defined as gain over rest, was an increasing function of amount of FH practice and length of rest between FH and SH practice.

It was suggested that an effector localization hypothesis of inhibition does not adequately account for the findings of bilateral reminiscence studies.

References

1. Ammons, R. B. Acquisition of motor skill: I. Quantitative analysis and theoretical formulation. *Psychol. Rev.*, 1947, **54**, 263–281.
2. Ammons, R. B. Acquisition of motor skill: II. Rotary pursuit performance before and after a single rest. *J. exp. Psychol.*, 1947, **37**, 393–411.
3. Ammons, R. B. Effects of pre-practice activities on rotary pursuit performance. *J. exp. Psychol.*, 1951, **41**, 187–191.
4. Ammons, R. B., & Ammons, C. H. Bilateral transfer of rotary pursuit skill. *Amer. Psychologist*, 1951, **6**, 294. (Abstract)
5. Bell, H. M. Rest pauses in motor learning as related to Snoddy's hypothesis of mental growth. *Psychol. Monogr.*, 1942, **54**, No. 1 (Whole No. 243).
6. Buxton, C. E. Level of mastery and reminiscence in pursuit learning. *J. exp. Psychol.*, 1943, **32**, 176–180.
7. Buxton, C. E. The status of research in reminiscence. *Psychol. Bull.*, 1943, **40**, 313–340.
8. Buxton, C. E., & Grant, D. A. Retroaction and gains in motor learning: II. Sex differences, and a further analysis of gains. *J. exp. Psychol.*, 1939, **25**, 198–208.
9. Cook, T. W. Studies in cross education: V. Theoretical. *Psychol. Rev.*, 1936, **43**, 149–178.
10. Eckstrand, G. A. Post-rest performance in motor learning as a function of amount of work performed in pre-rest practice. Unpublished doctor's dissertation, Northwestern Univer., 1949.
11. Ellis, D. S., Montgomery, V., & Underwood, B. J. Reminiscence in a manipulative task as a function of work-surface height, prerest practice, and interpolated rest. *J. exp. Psychol.*, 1952, **44**, 420–427.
12. Grice, G. R., & Reynolds, B. Effect of varying amounts of rest on conventional and bilateral transfer "reminiscence." *J. exp. Psychol.*, 1952, **44**, 247–252.
13. Helmick, J. S. Pursuit learning as affected by size of target and speed of rotation. *J. exp. Psychol.*, 1951, **41**, 126–138.
14. Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
15. Irion, A. L. Reminiscence in pursuit-rotor learning as a function of length of rest and amount of prerest practice. *J. exp. Psychol.*, 1949, **39**, 492–499.

16. Irion, A. L., & Gustafson, L. M. "Reminiscence" in bilateral transfer. *J. exp. Psychol.*, 1952, **43**, 321–323.
17. Kimble, G. A. An experimental test of a two-factor theory of inhibition. *J. exp. Psychol.*, 1949, **39**, 15–23.
18. Kimble, G. A. Performance and reminiscence in motor learning as a function of the degree of distribution of practice. *J. exp. Psychol.*, 1949, **39**, 500–510.
19. Kimble, G. A. Transfer of work inhibition in motor learning. *J. exp. Psychol.*, 1952, **43**, 391–392.
20. Kimble, G. A., & Horenstein, B. R. Reminiscence in motor learning as a function of length of interpolated rest. *J. exp. Psychol.*, 1948, **38**, 239–244.
21. Melton, A. W. The effect of rest pauses on the acquisition of the pursuitmeter habit. *Psychol. Bull.*, 1941, **38**, 719. (Abstract)
22. Snedecor, G. W. *Statistical methods.* (4th ed.) Ames, Iowa: Collegiate Press, 1946.

(Received February 22, 1953)

Journal of Experimental Psychology
Vol. 46, No. 5, 1953

TWENTY QUESTIONS: AN INFORMATION ANALYSIS

A. W. BENDIG[1]

University of Pittsburgh

A recent article by Taylor and Faust (2) reporting the use of a modified form of the parlor, radio, and television game "Twenty Questions" in studying concept formation suggested to the author that a similar experimental procedure might be useful in investigating the application of information theory to problems of human behavior. Generally the first problems investigated would be methodological: (*a*) the development of an experimental design amenable to information theory analysis, (*b*) the testing of the adequacy of specific procedures (stimuli, instructions, response measures, etc.), and (*c*) a check as to whether the design results in consistent changes in the experimental *S*s, changes presumably attributable to learning, fatigue, or variations in motivational level.

Two approaches were attempted to integrate information analysis with the game "Twenty Questions." The first was to modify the available mathematical procedures used in information theory to analyze the data provided by the traditional form of the game. Grave difficulties were immediately encountered because of the wide choice available to *S* as to which questions are asked and in what order, and in the smaller (but still considerable) number of answers that can be given by *E* to each question asked. This integrative attempt, after a valiant effort, was abandoned. Instead, the simpler Procrustean solution of modifying the game to fit the available analyses was adopted. First, a finite number of animal topics for the games was given to *S*s before and during each game and their responses were to be drawn only from this list. Second, the questions were asked and answered by the participating *E*s with the same questions in the same order being asked from game to game. Only the sequence of answers given by an *E* varied from game to game, depending on the topic selected. Only in this manner could we be sure that the same question would be asked in several games (yielding a replicated measure of the information in the question) and that each question would be asked in the same ordinal position within the game (holding constant sequential effects). With these modifications the game may no longer be the traditional form of "Twenty Questions," but it is a research tool capable of information analysis and of high motivational value for human *S*s.

Subjects and Procedure

The first step was the development of the questions and game topics to be used. From a series of questions that are commonly used at the beginning of the game the following four were selected on the rational basis of clarity and amenability to Yes-No answers:

1. Is the subject a whole animal?
2. Is the subject human?
3. Is there more than one of this subject?
4. Is the subject fictional?

With two alternative answers to each question (Yes-No) and four questions there are 16 possible permutations of answers. Several game topics were devised to fit each of these permutations and presented to graduate students for judgments as to whether the topics unequivocally corresponded to the Yes-No permutations. The 16 selected topics were then mimeographed on

[1] The author's grateful appreciation is extended to Messrs. James B. Hughes, James F. Horgan, and Stanley Klingensmith for assistance in collecting these data.

sheets with spaces for five subject responses after each topic. It might be noted at this point that the Yes or No answer to any one of the four questions rationally eliminated one-half of the topics from further consideration by the *Ss*.

The *Ss* were 127 daytime students enrolled in three sections of introductory psychology which were combined on the day that these data were collected. The *Ss* were told that the *Es* were studying the type of insight behavior often shown by panel members on the radio-TV program "Twenty Questions" and were interested in discovering whether college students could develop similar insight in this game. Sheets containing the 16 animal topics were given to the *Ss*, and they were informed that the topic for any one of the three games to be played could be found on this list. The topics had been singly typed on small cards and, after shuffling, one *S* randomly selected a card and handed it to one of the *Es*. Before the first question was asked *Ss* were asked to make a blind guess as to the topic selected and to record this guess on their mimeographed topic sheet. Then one *E* asked the first question (see above) which was answered by a second *E* as either "Yes" or "No." The *Ss* were then asked to record their second guess as to the correct topic. The second question was then asked and answered by the *Es* and the *Ss*' third guess recorded. This procedure was repeated for the remaining two questions. The *Ss* were informed of the correct topic after they had recorded their fifth guess.

Three successive games were played with the *Ss*, giving 15 successive responses from each *S*. For the second and third games the previously selected topics were replaced in the pack before a new topic was randomly drawn and *Ss* were informed of this. Parenthetically it may be noted that the overt motivation level of *Ss* remained high over the three games and increased ego involvement was apparent.

Analysis and Results

The first analysis of the data followed the information pattern detailed by Miller and Frick (1). Uncertainty (*U*) measures were computed from the proportions of *Ss* selecting each of the 16 topics on any single response. These are shown in Table 1. Theoretically, *U* should have had a value (in bits) of 4 on the first response, 3 on the second, 2 on the third, 1 on the fourth, and 0 on the fifth response. Obtained *U* was smaller than these rational values for Responses 1 through 3 and greater than these predicted values for Responses 4 and 5.

TABLE 1

Uncertainty (*U*) Measures for Each Response

Response	Game			Mean
	1	2	3	
1	3.61	3.67	3.53	3.60
2	2.62	2.82	2.87	2.77
3	1.83	1.80	1.95	1.86
4	1.62	1.57	1.77	1.65
5	0.73	0.77	1.11	0.87

The reduction of *U* between successive responses can be attributed to the information received by *Ss* by the asking and answering of one of the four questions between the pair of responses. Theoretically, each question should have reduced *U* by 1 bit, since one-half of the remaining topics were rationally eliminated by the answer. Actual information measures (*I*) for each question are given in Table 2. Approximately 85% of the information contributed by each question was transmitted by the *Ss*, except for Question 3. In a previous tryout of the questions on another group this question had been found to be the

TABLE 2

Information (*I*) Measures for Each of the Questions

Question	Game			Mean
	1	2	3	
1. "Is the subject a whole animal?"	0.99	0.85	0.66	0.83
2. "Is the subject human?"	0.79	1.02	0.92	0.91
3. "Is there more than one of this subject?"	0.21	0.23	0.18	0.21
4. "Is the subject fictional?"	0.89	0.80	0.66	0.78

most confusing, so special emphasis was placed on this question in the explanatory orientation given *Ss* prior to the games. In spite of the attempted clarification, only one-fifth of the information available in the question was transmitted to their responses by *Ss*.

A second analysis of the data was performed within a more conventional analysis of variance framework. On *S*'s second response, half of the 16 topics had been rationally eliminated by the answer to the first question. Choosing one of the eight remaining topics meant that *S* was still "on target," although only $N/8$ could be expected to be on the specific topic randomly selected. Being "off target" on the second response (selecting one of the eight categories rationally eliminated by the first question) could be considered a failure on the part of *S* to utilize fully the information in the question. Similarly, choosing one of the remaining four topics on the third response could be regarded as still "on target," while choosing one of the 12 eliminated topics (after the first two questions) was an "off-target" response. The 12 responses for each *S* (Responses 2 through 5 for Games 1 through 3) were scored as "on target" or "off target" and quantified by giving "on-target" responses a unitary weight of one and "off-target" responses a weight of zero. These quantified responses were analyzed as a factorial design by an analysis of variance and the summary can be found in Table 3. Differences between *Ss*, in addition to all of the single-order interactions, were significant at the 1% level when tested against the triple interaction (error). Differences between games were significant at the 5% level. Considering *Ss* as a random sample from the population, the between-responses mean square was tested against the significant responses-by-subjects mean square and was found to remain significant at the 1% level. Between-games was shown to be insignificant when compared to the significant games-by-subject interaction mean square. Percentages of *Ss* "on target" for each response in each game can be found in Table 4. The failure to transmit fully all the question information found in Table 2 is reflected in the increasing percentage of *Ss* going "off target" from Responses 2 through 4. However, a larger percentage of students were "on target" on Response 5 than were "on target" at Response 4. When compared by individual *t* tests, the mean differences between Games 1 and 2 and between Games 2 and 3 were not significant at the 5% level. The mean difference

TABLE 3

ANALYSIS OF VARIANCE OF "ON-TARGET" SCORES

Source of Variation	*df*	Mean Square	*F*
Games	2	0.2000	4.06*
Responses	3	2.5933	52.71**
Subjects	126	0.1691	3.44**
G × R	6	1.4300	29.07**
G × S	252	0.1317	2.68**
R × S	378	0.0773	1.57**
G × R × S	756	0.0492	
Total	1523		

* $p < .05$.
** $p < .01$.

TABLE 4

"ON-TARGET" PERCENTAGES ON EACH RESPONSE

Games	Responses				Mean
	2	3	4	5	
1	1.00	.95	.86	.87	.92
2	.95	.93	.80	.88	.90
3	.99	.96	.76	.82	.88
Mean	.99	.95	.81	.86	.90

between Responses 2 and 3 was significant at the 5% level, while "on-target" percentage differences between Responses 3 and 4 and 4 and 5 were significant at the 1% level.

Discussion

The major question as to whether the transmitting of information under the conditions of this study is subject to learning appears to be answered negatively. Although the between-games mean square in Table 3 is significant at the 5% level when compared to the triple-interaction (error) term, it drops to an insignificant level of confidence when tested against the significant games-by-subjects interaction mean square. No consistent trend in *U* measures from the first to the third game is evident. However, it must be noted that there is a confounding of the games and practice variables. A consistent decrease in *U* values and/or a consistent increase in "on-target" responses may be obscured by the selection of more difficult topics for the second and third games. Since the topic for any particular game was randomly selected from the available 16 topics such an explanation of the lack of significant game effects can only be attributed to chance. In the absence of other evidence the conclusion of no significant practice effect appears more parsimonious. The transmitting of information appeared to be more a function of the topic selected.

As noted previously, the 16 topics were rationally selected to fit the possible permutations of answers to the four questions. Wording of a few of the topics was defective. The topic "Kentucky Derby Winners" was used under the assumption that it referred to horses which have won the race in the past, but a number of *S*s evidently inferred that the reference was to humans (jockeys, horse owners, or bettors) who had won. Going "off target" on such a topic is less of a failure to assimilate information than it is a failure of the topic to unequivocally fit the category intended. Preliminary tryouts of topics are necessary.

The *U* and *I* measures in Tables 1 and 2 are remarkably consistent. *U* is always less than the predicted values on the first few responses, primarily because of *S*s' preferences among the topics on the first two responses. Chi-square tests of randomness of distribution of choices on the first response on each game yielded values significant beyond the 1% level. This, of course, reduced *U* from the expected value of four on the first response.

Regarding the questions used, Question 3 is obviously a poor one to use with these fairly naive *S*s. They are unable, at least without a lot of experience with this question, to utilize the information it contains. Future research should substitute another question, which would require a revision of the topics. Any possible sequential effect of the order in which the questions are asked was not testable with the experimental design used here.

Summary

An information analysis was made of the data collected from 127 *S*s participating in three games of a modified form of "Twenty Questions." The *S*s guessed animal topics from a list of 16 after each of four questions asked and answered by the *E*s. Results indicated that *S*s used 85% of the information provided by three of the questions, but transmitted only 21% of the information in a fourth question. No consistent increase in the transmitting of information attributable to learning was found over the three games although it was noted that the experimental design confounded practice and topic effects.

References

1. Miller, G. A., & Frick, F. C. Statistical behavioristics and sequences of responses. *Psychol. Rev.*, 1949, **56**, 311–324.
2. Taylor, D. W., & Faust, W. L. Twenty questions: efficiency in problem solving as a function of size of group. *J. exp. Psychol.*, 1952, **44**, 360–367.

(Received February 28, 1953)

Journal of Experimental Psychology
Vol. 46, No. 5, 1953

AN EFFECT OF REPEATED CONDITIONING–EXTINCTION UPON OPERANT STRENGTH

DONALD H. BULLOCK AND WILLIAM C. SMITH

University of Buffalo

The purpose of this experiment was to determine the effect upon the rate of an operant response of repeated reinforcement and nonreinforcement conditions. A number of previous studies have provided data relevant to the variable under investigation. Finch and Culler (3), Hilgard and Marquis (5), and Brogden, Lipman, and Culler (1) have all found that reconditioning of a reflex in an operant-respondent overlap situation requires fewer trials than the original learning. In addition, Brogden, Lipman, and Culler found that with repeated conditioning-extinction the number of reconditioning and re-extinction trials required to attain their criteria reduced progressively to a limiting value. Pavlov (6) has reported equivalent results for conditioned respondents. Shurrager and Culler (9) have obtained similar results in a spinal conditioning situation over four conditioning-extinction sessions.

With regard to "pure" operant conditioning, Ellson (2) found that a progressive reduction to zero in operant strength occurred with repeated extinctions without interpolated reconditionings. Wolfe (11), dealing with a single chimpanzee in a work situation, found a reduction to a limiting value of extinction responses per extinction period where conditioning and extinction sessions were alternated.

At the time of this research, however, no study had been reported in which the response rate during both reinforcement and nonreinforcement sessions was measured during repeated conditioning-extinction of an operant.[1] The experiment reported here was designed to provide initial information as to the effects of such a procedure on operant strength.

Method

The Columbia Skinner box and lever (4) were used, though modified in two ways: (*a*) the H slot for the lever was located to the rear of the cage on the left wall; and (*b*) the lever portion appearing inside the cage was covered with a metal sheet which reduced biting behavior to a minimum. The Skinner box was inside a large wooden icebox. The reinforcing stimuli, which were .05-gm., round, dry pellets of food, were manually delivered through a chute leading to the food tray. Water was available in the box at all times. The icebox was lighted inside by a red lamp. In addition, each lever press produced a 3-sec. duration illumination from a 25-w. frost bulb mounted at the top of the icebox.

The *S*s, six male albino rats of 90–110 days of age from the University of Buffalo colony, were placed on a 24-hr. feeding schedule one week prior to experimentation. The day before the conditioning-extinction sessions were begun, each rat was given a 1-hr. habituation session in the cage. Ten pellets were placed in the food tray. At the end of the hour, 40 light-food pairings were given to establish the light as an immediate secondary reinforcer. The light was manually controlled by *E*. During this training, behavior of turning away from the food tray and moving toward the lever location was reinforced in a series of progressive steps. The lever was absent at the time. On each of the next ten days *S*s were subjected to the following regimen: (*a*) their first 40 lever pressings were reinforced with single food pellets; and (*b*) starting with response 41, a 1-hr. period of nonreinforcement occurred, after which *S* was returned to his home cage. Three *S*s were run in the morning, three in the afternoon. They were food-deprived for approximately 20 hr. at the start of each session.

[1] The experiment by Perkins and Cacioppo (7) dealing with this problem was reported after completion of this study.

Results

Figure 1 presents the median cumulative response curves during the first 5 min. of the reinforcement portions of Sessions 1, 2, 3, 5, and 10. The response rate per minute for the first 5 min. of the ten sessions was: 1.0, 4.6, 6.2, 6.5, 7.3, 7.8, 7.2, 7.5, 7.1, and 7.8. These data indicate a stabilization of the rate beginning about the fifth day (approximately 200 reinforcements).

Figure 2 presents the median cumulative response curves during the 1-hr. extinction portions of Sessions 1, 5, and 10. The median total number of responses per hour for the ten sessions was: 52.5, 21, 30.5, 18, 27.5, 24.5, 13, 14.5, 13.5, 14.5. The median number of extinction responses per hour stabilized after about the sixth day (approximately 5–6 hr. of nonreinforced responding).[2]

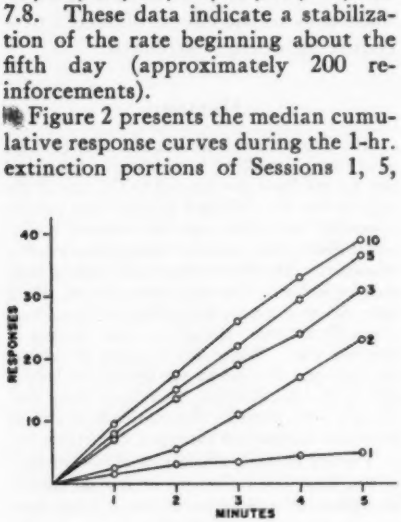


Fig. 1. The median cumulative response curves for the first 5 min. of the reinforcement portions of Sessions 1, 2, 3, 5, and 10

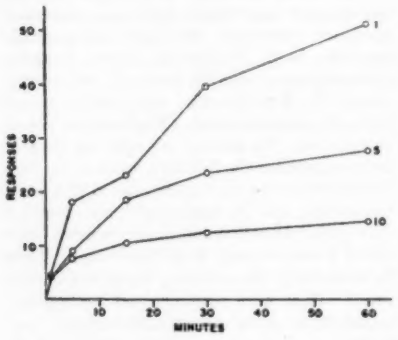


Fig. 2. The median cumulative response curves for the 60-min. extinction sessions of Sessions 1, 5, and 10

These data show that with repeated conditioning and extinction of the lever-pressing response, the rate during reinforcement conditions increased to a limiting value; and the rate during nonreinforcement conditions reduced to a limiting value.

Discussion

The results reported here bear an interesting similarity to standard discrimination data. Under the conditions of reinforcement, response strength increased and stabilized; under conditions of nonreinforcement, response strength decreased and stabilized. This similarity suggests an interpretation in line with the following hypotheses:

1. When a response is reinforced, those stimuli to which the organism is responding at the time gain selective control over the response. Such stimuli have been termed discriminative stimuli or S^D's by Skinner (10).

2. When a response is nonreinforced, those stimuli to which the organism is responding at that time lose selective control over the response through extinction. Such stimuli have been termed S^Δ's by Skinner (10).

3. Among the stimuli which can attain S^D or S^Δ functions are those stimuli (*a*) correlated with the occurrence of a previous response, or (*b*) directly produced by a preceding response.

[2] Raban reported data on repeated conditioning-extinction of lever pressing at the 1952 Eastern Psychological Association meetings. Though her procedure involved 20 reinforced pressings and 15 min. of nonreinforcement per session, her data attained limiting values after approximately the same total number of reinforced pressings and the same amount of nonreinforced time as that of this study.

4. Four contingencies can thus be established:

a. $R^1 \ldots S^R \ldots R^2 \ldots S^R$
b. $R^1 \ldots S^N \ldots R^2 \ldots S^R$
c. $R^1 \ldots S^N \ldots R^2 \ldots S^N$
d. $R^1 \ldots S^R \ldots R^2 \ldots S^N$

The S^N is used to designate nonreinforcement.

In the case of *a*, the sequence R^1–S^R would come to function as an S^D for R^2. Thus, if reinforced lever pressing is followed by reinforced lever pressing, the occurrence of reinforcement would exert discriminative control over lever-pressing behavior. With repeated occurrence of this contingency, the lever-pressing rate should increase to some limiting value, as has been demonstrated in this experiment.

In the case of *c*, the sequence R^1–S^N would come to function as an S^Δ for R^2. Thus, if nonreinforced lever pressing is followed by nonreinforced lever pressing, then pressing in the absence of reinforcement should progressively weaken through extinction, as this experiment has demonstrated.

Perkins and Cacioppo (7) have reported data which offer even more conclusive support to this interpretation. They compared extinction performances of rats given either 100% or 50% reinforcement during the reinforcement portions of repeated conditioning-extinction sessions. The 50% reinforcement group would be contingency *b* above, where nonreinforced pressing is followed by reinforced pressing, so that the absence of reinforcement becomes an S^D for pressing. This group showed a significantly lesser tendency to develop a discrimination between the reinforcement and nonreinforcement conditions than did the 100% reinforcement group.

Contingencies *b* and *c*, though not directly related to the data of the experiment reported here, are of general theoretical interest. Concerning *b*, where nonreinforced behavior acquires S^D control over a response, Sheffield (8) has demonstrated this result. Concerning *d*, where reinforced behavior should acquire S^Δ function, Skinner's (10) data on temporal discriminations developed under periodic reinforcement seem relevant. In such discriminations, the occurrence of reinforcement is followed by a significant reduction in rate. The inclusion within behavior theory of S^D and S^Δ functions of stimulation correlated with the occurrence of reinforced or nonreinforced behavior, whether these stimuli are external or internal, seems of particular significance for dealing with behavior traditionally referred to as "cognitive." Further research in this area should go a long way toward increasing the implications of reinforcement theory for so-called "complex" and "mediated" behavior.

Summary

Six male albino rats were subjected to ten successive daily conditioning-extinction sessions consisting of 40 reinforced lever pressings followed immediately by 60 min. of nonreinforced pressing.

Results showed (*a*) an increase in response rate to a limiting value under reinforcement conditions, and (*b*) a decrease in response rate to a limiting value under nonreinforcement conditions. The data are interpreted as exhibiting a discrimination based upon response-correlated stimulation.

References

1. Brogden, W. J., Lipman, E. A., & Culler, E. The role of incentive in conditioning and extinction. *Amer. J. Psychol.*, 1938, **51**, 109–117.
2. Ellson, D. G. Quantitative studies of the interaction of simple habits: I. Recovery from the specific and generalized effects of extinction. *J. exp. Psychol.*, 1938, **23**, 339–358.
3. Finch, G., & Culler, E. Relation of forgetting to experimental extinction. *Amer. J. Psychol.*, 1935, **47**, 656–662.
4. Frick, F. C., Schoenfeld, W. N., & Keller, F. S. Apparatus designed for introductory psychology at Columbia College. *Amer. J. Psychol.*, 1948, **61**, 409–414.
5. Hilgard, E. R., & Marquis, D. G. Acquisition, extinction, and retention of conditioned lid responses to light in dogs. *J. comp. Psychol.*, 1935, **19**, 29–58.
6. Pavlov, I. P. *Conditioned reflexes* (Trans. by G. V. Anrep) London: Oxford Univer. Press, 1927.

7. Perkins, C. C., Jr., & Cacioppo, A. J. The effect of intermittent reinforcement on the change in extinction rate following successive reconditionings. *J. exp. Psychol.*, 1950, **40**, 794–801.
8. Sheffield, V. F. Resistance to extinction as a function of distribution of extinction trials. *J. exp. Psychol.*, 1950, **40**, 303–313.
9. Shurrager, P. S., & Culler, E. Conditioning in the spinal dog. *J. exp. Psychol.*, 1940, **26**, 133–159.
10. Skinner, B. F. *The behavior of organisms.* New York: Appleton-Century, 1938.
11. Wolfe, J. B. Effectiveness of token-rewards for chimpanzees. *Comp. Psychol. Monogr.*, 1936, **12**, No. 5.

(Received March 2, 1953)

Journal of Experimental Psychology
Vol. 46, No. 5, 1953

THE EFFECT OF PHYSICAL CONSTANTS OF A CONTROL ON TRACKING PERFORMANCE [1, 2]

DANIEL HOWLAND AND MERRILL E. NOBLE
The Ohio State University

The two experiments described here were designed to test the general hypothesis that different combinations of control parameters will result in differences in performance in continuous compensatory pursuit tracking.

There have been several investigations of the effects of moment of inertia, gear ratio, and handwheel radius on tracking performance (**1, 5**), but no investigator seems to have studied the effects of combinations of spring stiffness, viscous damping, and moment of inertia. All three of these parameters determine the response characteristics of elastic systems when acted upon by external forces. This is true not only for mechanical systems, but also for electrical systems in which capacitance, resistance, and inductance are analogous to spring constant, viscous friction, and moment of inertia (**2**).

Different combinations of stiffness, damping, and inertia in a control may produce both quantitative and qualitative differences in performance because they may determine in part (*a*) the physical work required by a task and (*b*) quantitative and qualitative differences in the kinesthetic stimulation.

The nature of a tracking task may be a function of the distribution of work during the various phases of the task, and of the nature of the gradients of change of kinesthetic stimulation at different times during the movement pattern, as well as of the total work required by the task. These effects can be illustrated by the rotary back-and-forth motion of a control required to track a cursor performing simple harmonic motion. With reference to the pattern of work required by this task, both moment of inertia and spring stiffness exert their maximum effect at the ends of the cycles, where control motion is reversed. However, the effects of these two parameters are opposite, since moment of inertia increases the difficulty of reversing direction, while spring stiffness aids this reversal. The effect of viscous damping in opposing the rotation of the control is maximal at the center of the cycle, where rotational velocity is maximal, and decreases to zero at the ends of the cycle, where velocity is zero.

With reference to the kinesthetic stimulation produced during different phases of this movement cycle, position information is augmented primarily by spring stiffness, velocity information by viscous damping, and acceleration information by moment of inertia. Assuming that in the repetitive, predictable task provided by compensatory tracking of simple harmonic motion there is some sort of cyclical increase and decrease in kines-

[1] This research was supported in part by the United States Air Force under Contract No. AF 33(038)-10528 with the Ohio State University Research Foundation, monitored by the Perceptual and Motor Skills Research Laboratory of the Air Training Command, Human Resources Research Center. Permission is granted for reproduction, translation, publication, use and disposal in whole and in part by or for the United States Government.

[2] The writers are indebted to Dr. Paul M. Fitts who gave invaluable aid and advice throughout the progress of the study and, in collaboration with Mr. Paul K. Moffat, designed the hand control.

thetic information as a function of time, only the spring provides a steeper gradient of this information of a type that would be expected to aid *S* in anticipating when to reverse the motion of the control.

Method

Apparatus.—The single-dimension electronic pursuit apparatus employed in this study has been described previously (6). Essentially, it consisted of (*a*) an electronic apparatus which generated a single-dimension tracking problem made up of a stationary target and a cursor describing simple harmonic motion, (*b*) a cathode ray oscilloscope (CRT) on which the problem was displayed and on which *S* observed the error pattern, (*c*) a rotary hand control operated by *S*, (*d*) scoring and recording units, and (*e*) a programming unit.

The target was presented by generating a signal voltage V_p with the problem generator. This voltage was led to a subtractor. A second voltage, V_c, was generated by *S* as he rotated the control handle. This voltage was also led into the subtractor where the difference $V_p - V_c = V_e$ was formed. V_e was the error or difference between the output of the problem generator and that of *S*. This error was then fed from the subtractor to one terminal of a time-sharing switch which was connected to the plates of the CRT so that the horizontal position of the cursor was proportional to the magnitude of the error. The problem for *S* was to hold the cursor stationary below a target line.

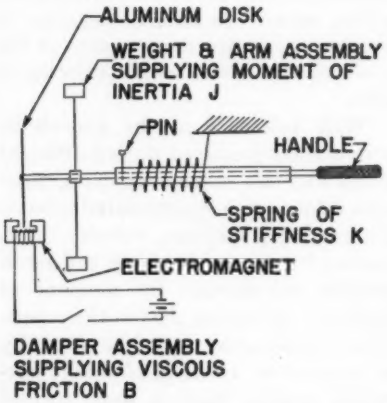


Fig. 1. Schematic diagram of the hand control

The display was presented on a 5-in. diameter CRT approximately 28 in. from the *S*'s eyes. When no problem was being generated, a 1-in. vertical line .5 mm. wide appeared in the center of the CRT. When the problem was being generated but *S*'s control was motionless, the upper half of the vertical line (the target) remained in place, while the lower half (the cursor) described simple harmonic motion below it with a frequency of 15 cycles per minute, and a total amplitude of 4 in. (approximately 15.96° of visual angle).

The control was a horizontal shaft which *S* gripped at one end in the manner that one grasps a screwdriver handle, and rotated through 40°.

The elements providing the various system constants were attached to the shaft as necessary to provide the various values of the independent variables required by the experimental conditions. The spring was mounted on a tube surrounding the hand-control shaft. This tube could be connected to the shaft by a pin. A collar, to which arms and weights were attached, supplied the moment of inertia effect when connected to the control shaft by a pin. Viscous-type damping was obtained by energizing a magnetic field in which an aluminum disk, geared to the control shaft, was rotated as the control handle was operated. Eddy currents set up in the disk opposed the rotation of the disk according to Lenz's Law. This resulted in a drag on the control handle which for any fixed strength of the magnetic field was proportional to the angular velocity of the disk. A 9-to-1 gear ratio was provided to increase disk speed for a given handle rotation. A schematic view of the hand control is shown in Fig. 1.

The computed values of the constants of spring stiffness (K), viscous friction (B), and moment of inertia (J), were approximately .10 ft.-pound/radian, .10 ft.-pound/radian/sec., and .01 ft.-pound/radian/sec.^{2}, respectively.

The characteristics of the time-varying torque, $f(t)$, applied by *S* to the handle of this apparatus can be expressed by the system equation

$$f(t) = K\theta + B\, d\theta/dt + J\, d^2\theta/dt^2$$

where K is the spring constant, B is the viscous friction, J is the moment of inertia, and θ is the angular displacement of the hand control.

Subjects.—A total of 80 right-handed men drawn from the elementary psychology courses at the Ohio State University served as the novice *S*s of Part I. Ten of the *S*s were assigned at random to each of the experimental conditions.

One practiced *S*, who was familiar with the experimental situation and who had had some previous practice on similar tasks, served in Part II.

Procedure.—Eight combinations of the control parameters were studied. These combinations are given in Table 1.

Compensatory tracking was studied because in this type of task only the error and its derivatives were visible. Since no visual information was available regarding the position of the control, it was assumed that *S*s would have to depend to a considerable extent upon proprioception for information while they were learning the task.

In Part I, each of the 80 *S*s tracked under one of the eight conditions for 20 1-min. trials separated from one another by 1-min. rest periods. During the rest period, the cursor remained stationary on the CRT. A warning buzzer sounded near the end of each rest period. In order to permit a short warm-up, cursor motion started several seconds before the time-on-target clocks began to record for the 1-min. scoring periods.

In Part II, the single *S* received three 1-min. trials per day, separated by 1-min. rest periods, on each of the eight conditions for 24 days, or a total of 72 trials per condition. Otherwise, the procedure was the same as that for Part I.

Indices of performance.—Two different time-on-target scores were recorded. These gave the total number of seconds during a 1-min. trial that *S* was able to hold the cursor within tolerances of ±15% and ±30% of the total amplitude of target oscillation. The absolute tolerance ranges, in terms of visual angle, were 2.45° and 4.90° for the two tolerance ranges.

Graphic records of *S*'s movement of the control handle as a function of time were also taken.

TABLE 1

OUTLINE OF EXPERIMENTAL CONDITIONS

Combination of Control Parameters	Symbol
"Free" control	OOO
Spring only	KOO
Viscous damping only	OBO
Moment of inertia only	OOJ
Spring and moment of inertia	KOJ
Spring and viscous damping	KBO
Viscous damping and moment of inertia	OBJ
Spring, viscous damping, and moment of inertia	KBJ

RESULTS

Part I.—The distributions of both the 15% and 30% tolerance scores were found to be somewhat positively

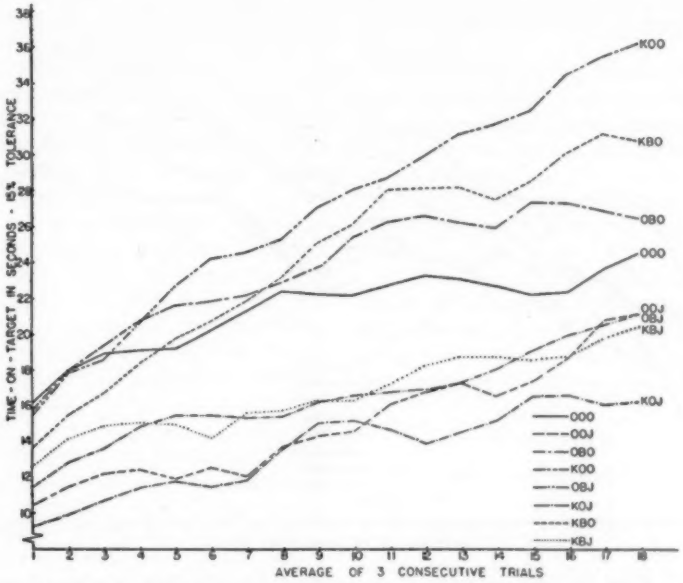


FIG. 2. Mean 15% tolerance scores of novice *S*s for the eight experimental conditions. Each point represents the mean of three adjacent 1-min. trials.

TABLE 2

ANALYSES OF VARIANCE AND *t* TESTS OF DATA IN PART I

Variance Source	*df*	Mean Square	*F*
F Tests of 15% Tolerance Scores			
Between conditions	7	261.54	4.21**
Within conditions	72	62.17	
F Tests of 30% Tolerance Scores			
Between conditions	7	416.08	3.73**
Within conditions	72	111.67	

Significant Individual t Tests of the 15% Tolerance Scores

KOO > KOJ**, OOJ**, OBJ**, KBJ**, OOO*
KBO > KOJ**, OOJ*, OBJ*
OBO > KOJ**, OOJ*
OOO > KOJ*

* $p < .05$.
** $p < .01$.

skewed, and the variances were not completely homogeneous. However, the χ^2's of Bartlett's test were significant only between the .05 and .01 levels of confidence, and the data were analyzed without transformation.

The mean 15% tolerance scores for the eight combinations of control parameters are shown in Fig. 2. The data for these curves were smoothed by averaging every three adjacent points. In general, the conditions having spring stiffness (K) result in the best performance, and those having moment of inertia (J) in the worst performance.

Analyses of variance indicated significant variation among the means of the various experimental conditions for both the 15% and 30% tolerance scores. Individual means for the 15% tolerance scores were then compared

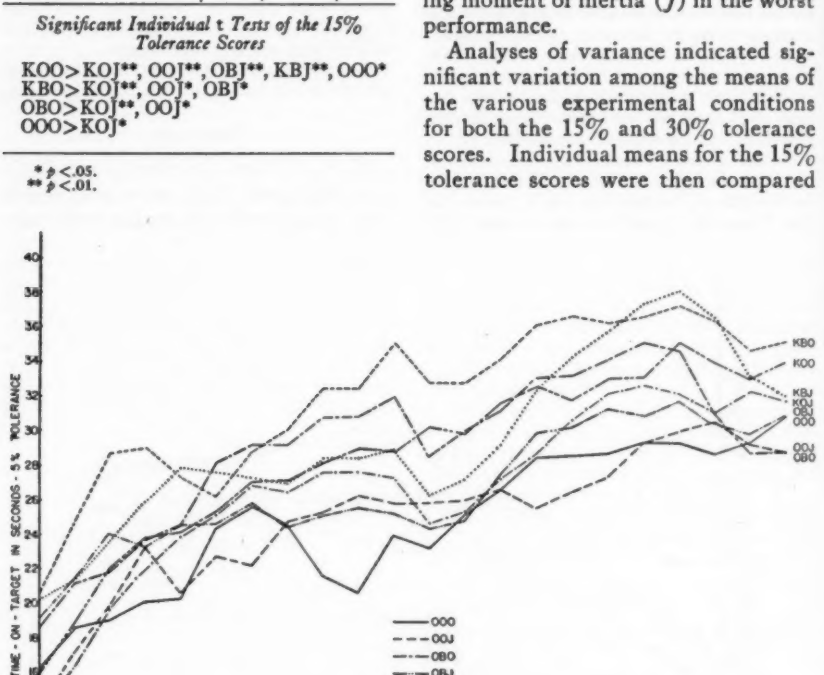


FIG. 3. Mean 5% tolerance scores of the practiced *S* for the eight experimental conditions. Each point represents the mean of nine 1-min. trials obtained during three adjacent days (three trials per day).

by t tests. Table 2 summarizes the results of the analyses of variance and also indicates which conditions differed significantly on the basis of the t tests (since all possible t tests were performed, the levels of confidence given in Table 2 are approximations).

Part II.—A 5% time-on-target score was recorded for the single practiced S. In Fig. 3 are shown the mean scores for the different combinations of control parameters, based on data for all 24 days. Again, in general, the combinations having spring stiffness (K) result in the best performance and those with moment of inertia (J) in the worst. An analysis of variance of these data indicated variation among means which was significant beyond the .001 level of confidence. Table 3 summarizes the analysis of variance and also shows which conditions were significantly different on the basis of t tests (again, p's are approximations).

Sample graphic records of the control movement of the practiced S at the start and at the end of the experiment are shown in Fig. 4.

Inspection of the graphic records of the single S at the beginning of practice indicates that differences in curve forms resulted from different combinations of control parameters. Under the "free" condition (OOO), there was a tendency for the peaks of the curves to be flat. This suggests that S reached the extreme right and left positions of the movement too early in the cycle and waited for the cursor. There also were irregularities throughout the cycle.

With the spring only (KOO), the curves became more saw-toothed than for the free condition, and continued to have irregularities. Viscous damping (OBO) resulted in even more saw-toothed curves with less irregularity in the middle of the cycles. This suggests a tendency to produce constant-rate movement patterns. Moment of inertia (OOJ) resulted in smooth curves except for a flattening at the control position corresponding to extreme pronation of the arm (bottom of the record).

TABLE 3

ANALYSIS OF VARIANCE AND t TESTS OF DATA IN PART II

Variance Source	*df*	Mean Square	*F*
F Test of 5% Tolerance Scores			
Days	23	217.55	34.42**
Conditions	7	145.70	23.05**
Error	161	6.32	

Significant t Tests of the 5% Tolerance Scores

KBO>OOO**, OOJ**, KOJ**, OBJ**, KOO**, OBO**, KBJ**
KBJ>OOO**, OOJ**, KOJ**, OBJ**
OBO>OOO**, OOJ**, KOJ**, OBJ**
KOO>OOO**, OOJ**, KOJ**, OBJ*
OBJ>OOO**, OOJ*
KOJ>OOO*, OOJ*

* $p < .05$.
** $p < .01$.

The various combinations of two or more of the control parameters produced smooth patterns, although for conditions having moment of inertia (J) as a component there was a tendency toward flattened peaks corresponding to the extreme right and left positions of the control.

At the end of practice (72 trials per condition) S was approximating a sine wave under all conditions, and no marked differences among the curves are apparent in the records.

DISCUSSION

The results are for the main part in agreement with the "principle of least work" advanced earlier on the basis of different distributions of work required by the different combinations of control parameters during various phases of the task. On such a basis the finding that

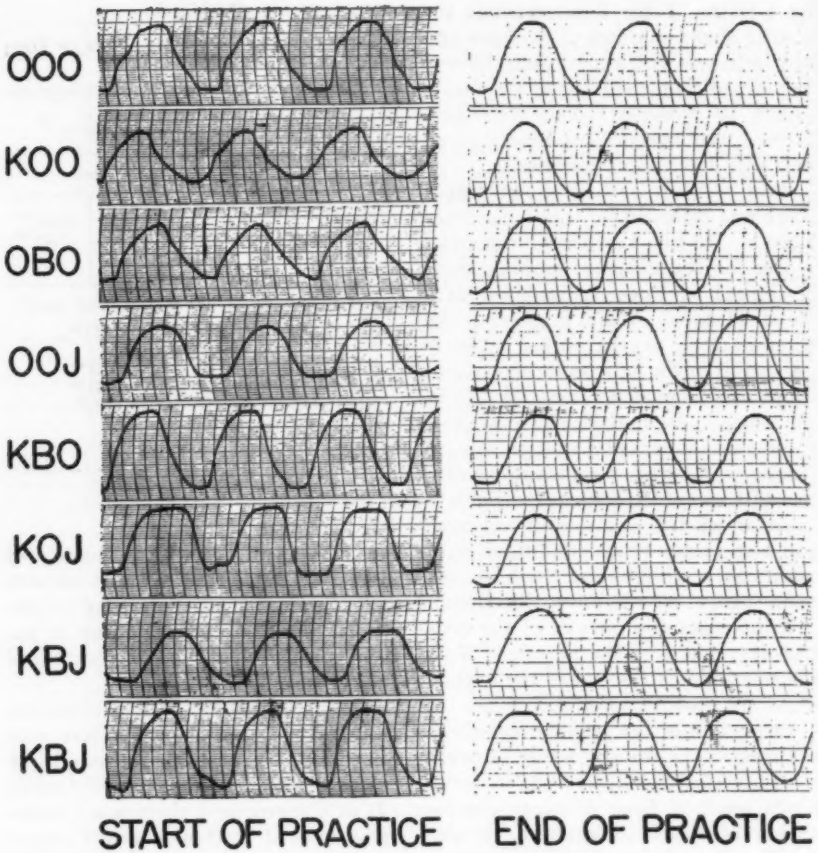


FIG. 4. Sample graphic records of the control movement of the practiced *S*. The records on the left were obtained at the start of practice and those on the right at the end of practice.

conditions that tend to aid reversal of the control movement (KOO and KBO) were superior to all other conditions would be expected. One exception to this explanation is that, in addition to KOO and KBO, the condition having viscous friction (OBO) resulted in better performance than the free condition (OOO).

Other exceptions were found in Part II. Interpretation of these is difficult, however, since all observations were made on a single *S* and thus the population to which statistical inference can be made is essentially undesignable unless the population is defined as all responses that *S* might have made. Since this is questionable, the *F*'s and *t*'s and the associated *p* values would be regarded primarily as descriptive rather than with the usual sampling interpretation.

An explanation in terms of least work, however, is not supported by data from certain other studies. For example, Helson (5) found that for a variety of types of target courses accuracy of compensatory tracking improved as handwheel speed and inertia increased, and

studies of following (3) and compensatory (4) tracking indicate that relative accuracy improves as the amplitude of control movement decreases.

As an alternative, or as an addition to the explanation in terms of different distributions of work requirements throughout the cycle produced by the different control parameters, the results can be accounted for by an explanation in terms of a hypothesis of differential feedback of kinesthetic information. From such a hypothesis it might be predicted that performance would be best for the combination of control parameters that maximizes the kinesthetic information most useful for this task.

The *S* probably makes use of several types of kinesthetic information in a tracking task—information about the position of the arm and the control handle, the rate of movement of the arm control, and the acceleration of these members. Variations in the control parameters increase the amount of stimulation of kinesthetic receptors that is normally present during movements. Variations in spring stiffness augment information about the displacement of the arm (θ), or position information. Velocity information ($d\theta/dt$) is augmented by the viscous damping. Acceleration information ($d^2\theta/dt^2$) is augmented by the moment of inertia of the control.

From the results obtained in these experiments it appears that augmentation of position information, supplied by the spring, is most helpful with the simple harmonic motion of the cursor studied here, with augmentation of velocity information next most helpful, and augmentation of acceleration information least helpful. The position information, augmented by the spring, may have been especially helpful in the compensatory tracking task because it requires less interpretation on the part of *S* than does velocity and acceleration information, and also because it assists in judging the proper amplitude of each successive movement.

Summary

The effects of various proprioceptive factors on learning and performance in a continuous compensatory tracking task were investigated by varying eight combinations of three system constants, a spring of stiffness *K*, a damping element of viscosity *B*, and moment of inertia of value *J*. The *S*s were required to apply a compensatory torque to a rotary control shaft so as to hold a cursor, which was describing simple harmonic motion, in alignment with a stationary target. The target and cursor were displayed as two vertical lines on the face of a 5-in. cathode ray tube. When they were in alignment they formed a single continuous line.

In Part I, eight experimental groups of ten novice *S*s each received 20 1-min. trials under a particular combination of system constants. In Part II, one practiced *S* was given three 1-min. trials per day on each of the eight conditions for 24 days. Time-on-target scores were taken for three different error tolerances.

Different system constants in the control resulted in significant differences in mean time-on-target performance. In general, performance was best for the conditions containing the spring and poorest for those containing the moment of inertia.

Variations in the control parameters influenced initial performance and these differences appeared to increase during the first 20 practice sessions in Part I. For the one *S* in Part II the effects appeared to decrease slightly as he became more skillful in the tracking task. Graphic recordings of hand-control position as a function of time were also taken. These records indicated distinctive movement patterns for the various conditions, particularly early in practice. These distinctive patterns tended to disappear with practice, so that the final records taken for the practiced *S* closely approximated the desired waveform regardless of the combination of parameters employed.

Explanations of these results in terms of amount of physical work and augmentation of proprioceptive factors were considered.

References

1. Craig, D. R., & Ellson, D. B. The design of controls. In *Human factors in undersea warfare*. Washington, D. C.: National Research Council, 1949. Pp. 133–151.
2. Fitts, P. M. Engineering psychology and equipment design. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 1287–1340.

3. Fitts, P. M., Marlowe, E., & Noble, M. E. The interrelation of task variables in a continuous pursuit task: I. Visual display scale, arm control scale, and target frequency in following pursuit tracking. USAF Human Resources Research Center, in press.
4. Hartman, B. O. An investigation of scale effects in simple, repetitive tracking tasks. Unpublished doctor's dissertation, Ohio State Univer., 1952.
5. Helson, H. Design of equipment and optimal human operation. *Amer. J. Psychol.*, 1949, **62**, 473–497.
6. Warren, C. E., Fontaine, A. B., & Clark, J. R. A two-dimensional electronic pursuit apparatus. USAF Human Resources Research Center, *Res. Bull.* 52–26, August, 1952.

(Received March 9, 1953)

Journal of Experimental Psychology
Vol. 46, No. 5, 1953

A QUANTITATIVE APPROACH TO FIGURAL "GOODNESS"[1]

JULIAN HOCHBERG AND EDWARD McALISTER

Cornell University

Empirical study of the Gestalt principles of perceptual organization is, despite their great heuristic value, frequently made difficult by their subjective and qualitative formulation. We wish to suggest here that it may be possible to achieve *parallels* to these "laws" of organization through analysis of the objective stimulus pattern. This approach differs from similar ones (**1, 6**) in the orienting hypothesis that, other things being equal, the probabilities of occurrence of alternative perceptual responses to a given stimulus (i.e., their "goodness") are inversely proportional to the amount of information required to define such alternatives differentially; i.e., *the less the amount of information needed to define a given organization as compared to the other alternatives, the more likely that the figure will be so perceived.*[2] However, to make this hypothesis meaningful, it is necessary to determine empirically the stimulus dimensions in which such "information" is to be measured. Therefore, we are concerned here mainly with so-called ambiguous stimuli (which evoke no single response with a probability of 1.0), although we will consider a possible theoretical bridge to the conditions of what Gibson and Waddell (**2**) call "determining stimuli."

An objective definition of perceptual "goodness" requires some measure of *Ss'* responses to stimulus figures. One such index might be the *threshold* (illumination, tachistoscopic, etc.), the "best" pattern having the lowest limen; however, this measure is too laborious for any really extensive survey of stimuli, and restricts the variety of stimuli which can be tested, being highly sensitive to recognition effects. Instead, we propose to use as a measure of "goodness" the response frequency or the relative span of time devoted by *S* to each of the possible perceptual responses which may be elicited by the same stimulus. This seems close to the intuitive meaning of "goodness" (**3**), and its probabilistic nature may permit rapprochement between perceptual laws on the one hand, and "information theory" (and, eventually, behavior theory) on the other (**5**).

That the concept of "information" (here meaning the number of different items we must be given, in order to specify or reproduce a given pattern or "figure," along some one or more dimensions which may be abstracted from that pattern, such as the number of different angles, number of different line segments of unequal length, etc.) may be useful in approximating figural "goodness" is suggested by almost any random selection of Gestalt demonstrations. The illusion of transparency obtains in Fig. 1a when less information is required to specify the pattern as *two* overlapping rectangles (number of different line seg-

[1] A slightly shorter version of this paper was read at the April 1953 meeting of the Eastern Psychological Association.

[2] After preparation of this paper, we have been privileged to see the manuscript of a paper by Dr. F. Attneave, which contains a much more detailed theoretical discussion of the tendency of the organism to perceive in terms of "maximum redundancy"; although this formulation is probably not precisely equivalent to the one proposed here, and the experimental techniques employed are quite different in method and assumptions, we are agreed as to the basic similarity of our general approaches.

ments: 8, plus one for notation of locus of intersection; number of angles to be specified: either 8, plus one for notation of angle of intersection or, more simply, one right angle plus the repetition implied in the notation of rectangularity, plus one for notation of angle of intersection; etc.) than, alternatively, as *five* irregular shapes (number of different line segments: 16; number of angles: 16; etc.). In Fig. 1b, less information is necessary to specify the symmetrical central black area as figure (number of different angles or points of inflection: 10, plus notation of duplication by bilateral symmetry) than the irregular white areas (number of different angles: 17). Listing the organizational "laws," from "good continuation" and "proximity" to the more general "simplicity" or "homogeneity," one finds translation impressively easy; the eventual utility of such translation depends, however, upon empirical determination of the dimensions of abstraction along which "information" is to be scored (shall we use "number of angles," "number of line segments," a weighted combination of these, or entirely different dimensions?), and upon the demonstration of a quantitative dependence of response frequency on the "information scores."

But can we approach the study of *determinate* (2) perception in this manner? Consider the task of representing spatial depth in two dimensions (Fig. 1c). Stimulus part *i* requires less specification as two overlapping rectangles than as L and rectangle; part *ii* requires less specification as a rectangle at a given slant (say 45°) than as a trapezoid at other slants; finally, in one way part *iii* requires less specification as three identical rectangles at the appropriate distances than as different sized rectangles at other distances (although here rather involved assumptions are necessary about a size-distance relationship "built in" to the specifying coordinate system, etc.). Now, although each part is ambiguous, if we take all the parts together and if the slant and depth relationships associated with the "best" response to each stimulus part should coincide, the probability of obtaining just these slants and depths will be reinforced. As we add more such "cues," the probability of obtaining alternative depth responses approaches zero, and we may therefore consider *determinate* perception as different from the ambiguous variety, with which we are concerned at present, not in kind, but in degree.

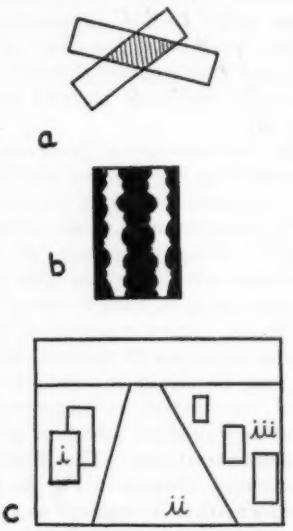


FIG. 1. Transparency, symmetry, and depth

METHOD

The approach outlined above is of little use unless it is possible to select dimensions for scoring "information" which are in correspondence with empirically obtained response-probabilities. The relative durations of alternate classes of response may be obtained by the usual method

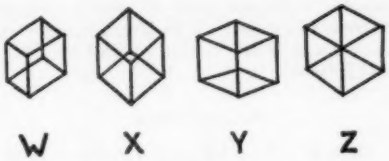


FIG. 2. The Kopfermann "cubes"

of pressing telegraph keys for each phase, but with this procedure *Ss* often report that the act of key-pressing altered the percept, which often fluctuated too rapidly to record; moreover, only one *S* can be used at a time. For these reasons, a sampling method was devised in which signal tones were presented by tape recording at "random" intervals and *Ss* indicated the phase they had perceived at the time each signal tone sounded. The frequency with which a given response is obtained is assumed to be proportional to the amount of time that response would have been obtained by "ideal" continuous recording.

The problem here was to apply this method to the case of Kopfermann cubes (Fig. 2) which may all be seen either as bidimensional patterned hexagons, or as tridimensional cubes (4). Drawings of each cube were presented in balanced order for 100 sec. each to 80 college students, providing a pool of over 2,600 responses for each stimulus; *Ss* indicated by pencil code marks which phase they had experienced just prior to each of the 33 signal tones presented at random intervals during the 100 sec.

RESULTS AND DISCUSSION

The results obtained correspond roughly to Kopfermann's more subjective findings: that figure which possesses the best phenomenal symmetry as a two-dimensional pattern was obtained least often as a cube (see Table 1). In terms of Gestalt theory, we would expect that the likelihood of seeing a figure in two dimensions would not only vary directly with its "goodness" in two dimensions, but, in addition, would vary inversely with its "goodness" in three dimensions. However, in this study, we may consider the "goodness" of the bidimensional patterns alone, since the tridimensional phase of each figure is more or less the same cube, the only appreciable difference being the apparent angle with respect to *S*. That is, we take the relative duration of two-dimensional responses to be proportional to the "goodness" of the two-dimensional patterns, the "goodness" of the tridimensional phases being approximately constant.

The bidimensional patterns may be analyzed for a large number of stimulus properties whose values will yield a relationship similar to that of the four points of bidimensional "goodness" response measures (Table 1), and the relationships fit quite well if these properties are differentially weighted. However, data are still needed on many other stimulus figures before a general system of factors and weights can be attempted, so that we are probably safer, at this stage, in merely noting stimulus variables which match the response relationships without employing differential weights. Two such stimulus dimensions fit the responses quite well, namely the number of angles and the number of line segments (Table 1). (Note that with the present figures, both dimensions represent the same geometrical fact.) Another dimension is the number of points of intersection required to define each bounded shape in the flat pat-

TABLE 1

BIDIMENSIONAL RESPONSES TO THE KOPFERMANN "CUBES" AND SOME TWO-DIMENSIONAL STIMULUS CHARACTERISTICS OF THE CUBES

"Cubes"	Bidimensional Responses (%)	Stimulus Characteristics		
		Line Segments	Angles	Points of Intersection
W	1.3	16	25	10
X	0.7	16	25	10
Y	49.0	13	19	17
Z	60.0	12	17	7

terns. Any of these scores would be consistent with an inverse relationship between response probability and the amount of "information," as discussed above, required to *specify* a given pattern; it is simple, however, to construct other figures whose relative strengths in alternate response phases may appear, at least intuitively, to be poorly handled by these dimensions, and we will need quantitative data from a wide sample of such figures before general stimulus dimensions can be chosen.

Summary

Probability of alternate perceptual responses is suggested as an approximate quantitative index of "goodness" of figure, and a group technique is presented by which this score can be obtained for ambiguous stimuli. Using the technique to obtain group scores for relative duration of tri- and bidimensional phases of four Kopfermann cube figures, the resulting responses are not inconsistent with the working hypothesis, namely that the probability of a given perceptual response to a stimulus is an inverse function of the amount of information required to define that pattern.

References

1. Brown, J. F., & Voth, A. C. The path of seen movement as a function of the vector-field. *Amer. J. Psychol.*, 1937, **49**, 543–563.
2. Gibson, J. J., & Waddell, D. Homogeneous retinal stimulation and visual perception. *Amer. J. Psychol.*, 1952, **65**, 263–270.
3. Koffka, K. *Principles of Gestalt psychology*. New York: Harcourt Brace, 1935.
4. Kopfermann, H. Psychologische Untersuchungen über die Wirkung zweidimensionaler Darstellungen körperlicher Gebilde. *Psychol. Forsch.*, 1930, **13**, 293–364.
5. Miller, G. A. *Language and communication*. New York: McGraw-Hill, 1951.
6. Orbison, W. D. Shape as a function of the vector-field. *Amer. J. Psychol.*, 1939, **52**, 31–45.

(Received for early publication May 21, 1953)

Journal of Experimental Psychology
Vol. 46, No. 5, 1953

SIMULTANEOUS VISION AND AUDITION: THE COMPREHENSION OF PROSE PASSAGES WITH VARYING LEVELS OF DIFFICULTY [1]

G. H. MOWBRAY
The Johns Hopkins University

Recent interest in the development of visual and auditory communication systems has led to a consideration of the relative efficiency of vision and audition under conditions of simultaneous stimulation. It has often been shown that a combined visual and auditory presentation of the same material results in better comprehension than the presentation of either auditory or visual material alone (**2**). Very little is known, however, about the relative performances of these two modalities when they are simultaneously stimulated by conflicting material. Recently, Mowbray (**10**) has shown that for combined tasks involving more than one level of difficulty, significant deterioration of performance, compared to that with nonsimultaneous presentation, caused by the conflicting nature of the material occurred with the easier of the tasks. No differential modality effects were detected. That is to say, deterioration of performance from nonsimultaneous to simultaneous trials was not significantly greater for one modality than the other. The significant deteriorations that did occur were attributable to the level of difficulty of the experimental material. When a relatively easy task was paired with a relatively difficult one, the easier task suffered the most. The imposed task required Ss to detect and report elements randomly deleted from alphabet and numeral sequences. Presentation of visual and auditory sequences was such as to permit rapid alternations of attention between the two sensory channels.

The importance of this result is limited by the somewhat unusual nature of the perceptual task employed. Therefore, it was thought that by providing perceptual situations of a more familiar kind it might be possible to extend or restrict the generalization of this finding.

Considerable theoretical interest is attached to the problem of the unitary as opposed to the multiple nature of attention. Woodworth (**12**) has reviewed some of the classical experiments dealing with this question. Most of the investigations he described were concerned with a combination of a sensory and a motor task, and in almost all cases some impairment of performance occurred with either one or both of the tasks. The only research that utilized simultaneous sensory tasks forced the conclusion that simultaneous performance of two attentive

[1] This research was partially supported and done in cooperation with the Systems Division, Naval Research Laboratory under Contract N5-ori-166, Task Order I, between the Office of Naval Research and The Johns Hopkins University. This is Report No. 166-I-172, Project Designation No. NR 507–470, under that contract. The actual experimental work was done at the Psychological Laboratory, University of Cambridge, Cambridge, England, and the author is grateful to Professor Sir Frederic Bartlett and his successor, Professor O. L. Zangwill, for the facilities and assistance provided at that laboratory. Dr. N. H. Mackworth was particularly helpful in reading and criticizing the completed manuscript.

acts of cognition seldom, if ever, occurred (**12**, p. 712).

Hebb (**6**), in a current attempt to structure the field of perceptual theory, has touched briefly on the problem of division of attention. His hypothesis implies a denial of the unity of attention and, in fact, he explicitly asserts that there are no empirical data to controvert such an implication (**6**, pp. 151–152). He states that more than one phase sequence may be in operation in the association areas at the same time, presumably leading to simultaneous perceptions in separate sense modalities. Only anecdotal evidence is cited to support this contention, but that, coupled with his phase-sequence hypothesis, leads him to suggest that attention may often be of a multiple nature.

With these considerations in mind, the purpose of this experiment was twofold. First, it was thought that by providing visual and auditory tasks that were familiar and meaningful, further information might be forthcoming as regards the interaction of difficulty of stimulus material and sensory channels as reported earlier by Mowbray. Second, by restricting or removing completely the possibility of rapid alternations of attention between modalities, data might be furnished on the problem of the unity of attention. Thus it might be possible to provide some data that would either strengthen or weaken Hebb's position on this question.

Accordingly, an experiment was designed that utilized prose passages of varying levels of difficulty drawn from fictional sources.

Method

Materials.—The prose passages were selected from a wide range of fictional sources that were judged to display different levels of difficulty. They were chosen for interest of content, general unfamiliarity of the work, and length, since presentation of visual and auditory extracts was to be simultaneous. When a passage was judged to fall into one of the categories of "difficult," "medium," or "easy," from the standpoint of comprehensibility, it was subjected to the Flesch readability formula (**3**). This formula, when applied, provides a rating on a scale of 0–100, where 0 represents extreme difficulty and 100 represents material that is comprehensible to persons defined as "barely, functionally literate." If passages judged to be "difficult" provided ratings from 0–30, they were selected; "medium" passages falling in the range of 31–60, and "easy" passages within the range 61–100 were likewise chosen. In all, a total of 54 selections were made—18 for each level of difficulty defined. The mean reading-ease scores for the difficult, medium, and easy passages were 21.6, 46.7, and 72.3, respectively, and the corresponding ranges were 5–20, 39–58, and 61–82. The average length of the passages was 179 words, but they ranged from 154 to 220 words. In many cases the passages were altered to prevent recognition of the sources from which they were gathered. Alterations were generally restricted to changing the names of persons and places.

Simultaneous pairings of the visual and auditory test material permitted nine possible combinations of the three levels of difficulty. These pairs were repeated three times, and half of them were recorded by a female voice on a magnetic tape recorder. The other half for visual presentation were typed on a 6-in. wide roll of ordinary paper recording tape. The various pairs of combinations of levels of difficulty were distributed randomly throughout the series. This constituted Series I. Series II was exactly the same with the exception that the passages that were visual in Series I were auditory in Series II; and those that had been auditory were, of course, visual in this latter case.

The apparatus for visual presentation of the passages was designed to pace the reading of Ss to match the speed of the aural stimuli and to restrict forward and backward glancing. A constant speed of approximately 175 words per minute was maintained throughout the course of the experiments. A thyratron-driven motor pulled the tape, on which the visual passages were typed, past a viewing window cut in a gray metal screen. The viewing window was the same width as one line of type, and it was long enough to allow vision of the entire 6 in. of the recording tape. However, the viewing window was so tilted that relatively normal eye movements could be maintained during reading, but forward and backward glancing was restricted to about one-third of a full line.

Synchronization of visual and auditory material was adequately maintained by careful timing of the magnetic tape recorder and the visual projector.

Procedure.—The *Ss* were seated in a quiet room facing the visual viewing apparatus described above. The magnetic tape recorder was placed to their left and about 3 ft. away in a direct line with their ear. They were told that they were to be presented with two different prose passages at the same time; one they would have to read for themselves and the other they would have to listen to. They were asked to try to understand as much of both selections as they could because they were going to have to answer questions about both at the end of the presentation. There were ten questions of a true-false nature for each paragraph. They were designed to test for general understanding of the paragraph as well as for the perception of selected facts, with the greatest emphasis on the factual side. The *Ss* were asked to consider the questions only in the light of what they had just heard or read and to remember that the sources were fictional and therefore had license to wander from the truth as they might understand it from personal experience. If they did not know the answers to any questions, they were asked to guess.

After each pair of simultaneous presentations, *Ss* were given the appropriate ten questions in an alternating auditory-visual, visual-auditory order that was consistently maintained for the course of the experiment.

Two 1-hr. experimental sessions were devoted by each *S* to completion of the 27 pairs of simultaneous presentations.

Upon completion of the experiment, all *Ss* were given a list of books and authors that included those from which the experimental material had been selected. They were asked to state which of the books they had read. Those that were among the experimental list were noted, and *Ss* were asked if they could identify any of the passages. Any identifications were likewise recorded.

Control *Ss* were presented with the same material as the experimental *Ss*, but were required to use only one modality at a time. The order in which the various passages were presented followed the question-order of the experimental *Ss*, i.e., auditory-visual, visual-auditory, etc.

Three 40-min. sessions were devoted to completion of the 54 auditory and visual presentations in the case of control *Ss*.

Subjects.—There were two distinct groups of *Ss* providing two separate but parallel experiments. One group was composed of university undergraduates, research workers, and graduate students from various laboratories, and a few others participating to a great extent in university life. The other group was composed entirely of enlisted men from the British Navy who were selected on the basis of performance on the AH4 Intelligence Test (7). Only those enlisted men whose scores fell in the 90th percentile, according to established norms, were used.

Of the nonnaval *Ss*, there were 28 experimental and 12 control *Ss* who were between 19 and 36 years of age. There were 27 males and 13 females. The naval *Ss* consisted of 30 males—20 experimental and 10 control—whose ages ranged from 18 to 25. With both groups of *Ss*, one-half of the experimental group and one-half of the control group were assigned at random to Series I and the other half to Series II.

In view of the forced-guessing techniques used and the resulting likelihood of obtaining 50% correct answers purely by chance, some check was required on a possible bias written into the questions that might permit a better than chance response. Accordingly, three naval and three nonnaval *Ss* were required to give answers to the 540 questions without the benefit of ever having seen the material on which the questions were based.

Results

Upon analysis it was clear that the results from the two groups of *Ss* did not differ in any essential way. Therefore, all reported results will consider the two groups as one.

Table 1 gives the average percentage of correct responses for the three levels of difficulty of passages and for visual and aural presentation. The results of the control *Ss* show how the

TABLE 1

Percentage of Correct Responses on Tests of Retention for Prose Passages Presented Visually, Aurally, or Simultaneously Visually and Aurally

Difficulty of Prose Passage	Control Group		Experimental Group*	
	Visual	Auditory	Visual	Auditory
Difficult	60.8	63.3	58.2	56.5
Medium	67.4	67.3	58.1	56.1
Easy	74.0	73.1	63.4	59.5

* Passages presented visually and aurally simultaneously.

obtained comprehension scores bear out the objective rating of the material. A test for the significance of the difference between the difficult and the medium categories and between the medium and the easy shows significance in the case of all visual trials ($p < .001$). The auditory results show the same trend with the exception that the difference between the difficult and the medium categories cannot be accepted with as much confidence ($p < .02$). The difference between the medium and easy is, however, highly significant ($p < .001$).

The scores of the control *Ss* who were required to answer all of the questions without benefit of having seen the material from which the questions were drawn did not differ significantly from chance, so it may be concluded that the questions, per se, introduced no positive or negative bias in the answers.

As was to be expected with simultaneous presentation of visual and auditory materials, in every case the performance level dropped from what would have been anticipated on the basis of nonsimultaneous reception. On the average, none of the categories dropped as low as chance performance, although the difficult and the medium material presented aurally came close to it. However, all scores in Table 1 are significantly greater than would be expected on a chance basis ($p < .001$).

TABLE 2

AVERAGE PERCENTAGE DETERIORATION OF CORRECT RESPONSES DUE TO SIMULTANEOUS PRESENTATION

Modality	Level of Difficulty		
	Difficult	Medium	Easy
Visual	2.6	9.3	10.6
Auditory	6.8	11.2	13.6

Table 2 makes the comparisons between the performance of the experimental and control *Ss* for the same materials. The scores represent the average percentage of deterioration in response due to the simultaneous pairing of the passages. They were obtained from Table 1 by subtracting the scores for experimental *Ss* from the appropriate scores for control *Ss*. Two things of importance are to be noted in these deterioration scores. First is the fact that in all cases the percentage of deterioration of performance is increasingly greater from difficult to medium and from medium to easy. This is true with both the visual and the auditory material. The second is that the deterioration is greater for the auditory material than for the visual.

In order to test more rigorously the significance of these two trends, an analysis of variance was performed. The scores for analysis were obtained by averaging the nine scores for each *S* on each condition of level of difficulty and sensory channel stimulated. These scores for the individual *Ss* of the experimental group were subtracted from the average score of the total control group for the appropriate condition.

In view of the fact that all *Ss* of the experimental group performed all conditions of the experiment, a method of extracting variances suggested by McNemar (**9**, p. 306) was utilized.

Table 3 presents the results from the analysis of variance. In the case of both the channel stimulated and the level of material difficulty, the variances are significant ($p < .001$ in both instances). That is to say, significantly more deterioration occurred with the auditory material than with the visual, and further, the easier material suffered relatively the most with the combined visual and auditory pres-

entation. Also, the variance contributed by Ss shows a significant departure from chance ($p < .001$).

The results already cited seem to indicate that some measure of simultaneous reception of the prose passages did occur. Before anything decisive may be concluded about that, however, it is necessary to look at all simultaneous pairs to see if, in the aggregate, the lowest scores of each pair, regardless of channel, are significantly greater than chance. The results of such a pair-by-pair analysis are given in Table 4. In this table, the low scores were obtained by averaging the lowest of the visual or auditory scores obtained from each presentation of each condition to each S. The high scores are a similar average. In every case except the combination of easy passages with easy passages it is evident that response to one or the other of the paired presentations was made at a chance level; and, in truth, the significance of the departure from chance of the lowest scores of the easy-easy combination is in doubt ($p < .05$).

A final consideration concerns the degree of familiarity, if any, that Ss had with the works of fiction from which the experimental materials were chosen. Of the naval Ss, only three of the controls and none of the experimental group confessed to having read any of the books involved. The nonnaval group, as might be expected, reflected a somewhat wider literary experience. Of these, 6 of the control group and 11 of the experimental group expressed some familiarity with one or more of the novels. However, only two Ss, both nonnaval, could connect any passage with its proper source, and they both admitted that their previous experience provided no aid in answering the questions. The truth of such an admission must be accepted on faith, but, in any case, the number of passages involved is so small as to be of almost no importance in the over-all picture.

TABLE 3*

ANALYSIS OF VARIANCE OF DETERIORATION SCORES OBTAINED FROM A COMPARISON OF SIMULTANEOUS WITH NONSIMULTANEOUS PERFORMANCE
$N = 48$

Source	df	Mean Square	F
Channel	1	58,910.28	18.07**
Material	2	162,728.64	49.93**
Subjects	47	9,916.27	3.04**
C × M	2	2,093.90	
C × S	47	3,198.79	
M × S	94	2,377.08	
C × M × S	94	3,259.32	
Total	287		

* Scores used for analysis were the average deterioration scores + 1.00 × 100.
** $p < .001$.

TABLE 4

AVERAGE PERCENTAGE OF CORRECT RESPONSES OF LOWEST SCORE AND HIGHEST SCORE FOR ALL SIMULTANEOUS PAIRS

Simultaneous Pairs	High	Low
Difficult—difficult	66.6	49.9
Medium—medium	65.9	48.8
Easy—easy	71.9	52.5*
Difficult—medium	68.8	49.9
Difficult—easy	68.8	50.7
Medium—easy	64.8	48.8

* $p < .05$.

DISCUSSION

Previous research on the efficacy of visual vs. aural presentation of meaningful, connected prose to adults provides results that are in conflict. With material of approximately the same degree of difficulty, some investigators have claimed the visual mode to be superior for immediate recall (1, 11), while others have found aural supremacy or else no difference between the two modes (5, 13). When the levels of material difficulty are varied, it has been reported that visual

presentation is better for material of moderate difficulty, but that there is no difference with easier material (8). Somewhat the reverse state of affairs has been reported by Goldstein (4), who found that difficult material was handled better aurally, but that the difference between visual and auditory methods of presentation became less as the material became easier. The data from the controls in this experiment tend to support the view that aural presentation may be better for difficult material (although the difference was not significant), and that there is very likely no difference at all with easier material. One should indeed be extremely wary of drawing such a conclusion, however, in view of the somewhat unnatural method of visual presentation employed in this research. Most of the investigations cited above made use of normal reading habits. In such cases very little control can be exercised over the amount of time spent on various portions of the stimulus material. Difficult and involved elements may be referred to more than once, and some advantage may thus be gained over aural presentation, which does not present similar opportunities. This may explain the reported advantage of visual presentation for relatively difficult material. The results of this experiment suggest that when such an advantage is removed the difference between modalities tends to disappear.

Of much more immediate interest are the results from the simultaneous trials. The findings, as indicated above, show the same consistent trend that was reported by Mowbray (10), namely, the significantly greater vulnerability of the easy material when a conflict is established between stimuli that represent different basic levels of difficulty. Neither result, alone, provides much ground for confidence that the effect is anything more than an experimental artifact. Each research has critical shortcomings; the first because of the unfamiliar and, in a sense, unrealistic nature of the perceptual task, and the second because of the performance limit imposed by the operation of the laws of probability. However, the fact remains that two vastly different perceptual situations have yielded essentially similar results.

One likely suggestion to explain this result in the latter experiment concerns the nature of the experimental material. The difficult prose passages were difficult for several reasons—grammatical style, unusualness of vocabulary, length of sentence, and abstract character of content. Conversely, the easy selections were characterized by short sentences, common words, and marked specificity of content. They were full of discrete facts and depended on those facts to make their theme coherent. The questions for each type of material were naturally influenced by these considerations. It seems very likely that the general redundancy of prose is accentuated with the more abstract themes and that a good many more words might be lost in the abstract selections than in the factual ones with correspondingly less loss of understanding. There is no available evidence for this, and it is offered here merely as a possibility that merits further attention.

Where Mowbray (10) found no significant differences in deterioration between the two sensory modes, this research has shown a relatively greater tendency for the auditory channel to suffer under the simultaneous conditions. The manner in which the visual material was presented may have been responsible for this result. It is a natural state of affairs to have to do nothing except listen to an auditory flow of information, but it is distinctly unnatural to have one's reading paced by an external source. This departure from the usual order of things compelled *Ss* to attend constantly to the ever-moving visual stimulus or risk losing large and important sections of it. That was the object of the experiment, but this imposed compulsion to *do* something was bound to affect the auditory channel unless complete division of attention could be maintained.

The evidence provided by this research on the possibility of successful division of attention, while not entirely conclusive, is, at the very least, compelling. There is reason to believe that for nearly

all pairs of simultaneous combinations, one or the other of the pair was responded to on a chance basis. Some doubt must remain, however, since one of the combinations (the easy with easy) yielded a result that suggests some measure of success in dividing the attention between two simultaneously presented sources of stimulation.

In spite of the limitations placed on backward and forward glancing with the visual trials, *Ss* were still able to accomplish some measure of alternation of attention. The visual material visible at any one moment (about one-third of a 6-in. line) could be assimilated in a glance with a short pause for attention to the auditory channel before enough of the line had moved on to make a further glance necessary. Continuing in this manner, some approach to successful accomplishment of the task was probably made possible. Nearly all *Ss*, according to their report, consciously adopted this procedure during the course of the experiments. Many were firm in declaring that if some such compromise were not possible, the task would have been beyond their capabilities.

The suggestion then, on the basis of the experiment described, is that as alternation of attention becomes increasingly more difficult, the possibility of performing adequately both of two simultaneous perceptual tasks becomes more remote. This evidence tends to support the thesis that attention is unitary, at least with respect to complex, cognitive acts, and is in harmony with previous findings. Before it is possible to state conclusively that Hebb's position on this matter is unwarranted by the facts, more data are required on a wider range of stimulus materials.

Summary and Conclusions

Prose passages selected from fiction sources and having three levels of difficulty as determined by a Flesch count were presented to *Ss* visually and aurally at the same time. Two groups of *Ss* were used—one a group of naval enlisted men and the other a group from a university population.

The auditory material was recorded on a magnetic tape recorder, whereas visual presentation was accomplished by means of a visual pacer that prevented forward and backward glancing. The speed of visual and auditory presentations was identical.

Response was measured by a ten-item questionnaire designed to test for comprehension of the paragraph and for the retention of discrete facts. Scores obtained from simultaneous trials were compared with scores obtained from control groups that performed the same tasks nonsimultaneously.

The following conclusions can be drawn:

1. For simultaneous presentation significantly greater deterioration (from a nonsimultaneous level) occurred with the easy material, thus substantiating earlier findings.

2. When scores from all paired presentations were separated on the basis of the lowest score and the highest score regardless of channel involved or level of difficulty of material, it was found that the lowest scores, with only one exception, did not differ significantly from a chance level. It was suggested that when the opportunity for rapid alternation of attention between sensory channels is severely reduced, the possibility of adequately performing both of two simultaneous perceptual tasks becomes more remote.

3. The auditory channel showed significantly greater disruption of performance than the visual for the simultaneous operation.

References

1. Corey, S. M. Learning from lectures versus learning from readings. *J. educ. Psychol.*, 1934, **25**, 459–470.
2. Day, W. F., & Beach, B. R. A survey of the research literature comparing the visual and auditory presentation of information. Air Materiel Command, United States Air Force, Technical Report No. 5921, 1950.
3. Flesch, R. A new readability yardstick. *J. appl. Psychol.*, 1948, **32**, 221–233.
4. Goldstein, H. Reading and listening comprehension at various controlled rates. *Teachers Coll. Contr. Educ.*, No. 821. New York: Bureau of Publications, Teachers Coll., Columbia Univer., 1940.
5. Green, E. B. The relative effectiveness of lecture and individual reading as methods of college training. *Genet. Psychol. Monogr.*, 1928, **4**, 457–563.
6. Hebb, D. O. *The organization of behavior*. New York: Wiley, 1949.
7. Heim, A. W., & Wallace, J. G. The effects of repeatedly retesting the same group on

the same intelligence test. Part I. Normal adults. *Quart. J. exp. Psychol.*, 1949, **1**, 151–159.

8. LARSEN, R. P., & FEDER, D. D. Common and differential factors in reading and hearing comprehension. *J. educ. Psychol.*, 1940, **31**, 241–252.
9. MCNEMAR, Q. *Psychological statistics*. New York: Wiley, 1949.
10. MOWBRAY, G. H. Simultaneous vision and audition: the detection of elements missing from overlearned sequences. *J. exp. Psychol.*, 1952, **44**, 292–300.
11. RUSSELL, R. D. A comparison of two methods of learning. *J. educ. Res.*, 1928, **18**, 235–238.
12. WOODWORTH, R. S. *Experimental psychology*. New York: Holt, 1938.
13. WORCESTER, D. A. Memory by visual and by auditory presentation. *J. educ. Psychol.*, 1925, **16**, 18–27.

(Received for early publication July 10, 1953)

Journal of Experimental Psychology
Vol. 46, No. 5, 1953

AN INFORMATIONAL ANALYSIS OF ABSOLUTE JUDGMENTS OF LOUDNESS [1]

W. R. GARNER

The Johns Hopkins University

In a paper by Garner and Hake (2) it was shown that information measurement is a potentially useful way of studying problems of perception and judgment, and several recent papers (3, 4, 8) have demonstrated the experimental feasibility of this approach. Of particular interest for the topic of the present paper is the paper by Pollack (8), which is concerned with the absolute judgment of pitch.

A measure of information transmission provides a means of specifying perceptual or judgmental accuracy in situations where absolute judgments about various categories on a stimulus continuum are required. This measurement allows the determination of the maximum number of stimulus categories which could be used with perfect accuracy without the necessity of sampling all the possible numbers of categories. However, this use of information transmission requires the assumption that the inherent judgmental accuracy is independent of the number of stimulus categories used experimentally. Two experiments (3, 4) have shown that this assumption is quite valid for situations involving judgments of position in visual space, and Pollack's experiment demonstrates its validity for judgments of pitch.

The purpose of the present experiment is (*a*) to determine the effect of number of stimulus categories on accuracy of judgments of loudness and (*b*) to examine some of the factors which affect judgmental accuracy in this type of experimental situation.

METHOD

Observers.—Six *O*s were used in this experiment, two female and four male, ranging in age from 20 to 30 years. The number of *O*s was limited because complete data on individual *O*s were required for the analysis.

Procedure.—One *O* at a time was seated in a soundproof room and was given a pair of earphones wired for monaural listening. He listened to a series of tones and was required to write on a sheet of paper a number which was his judgment about which tone he had just heard. The numbers used varied from experiment to experiment, but *O* was always instructed to use as many numbers as there were stimulus categories for the particular experiment. A particular series of tones consisted of 100 judgments, after which a new series was started. Each experiment was completed before another began, but the various experiments were run in counterbalanced order for the different *O*s. Experimental sessions lasted for approximately 2 hr. each, with *O* being allowed rest periods as often as he liked.

The tones heard by *O* were of 1-sec. duration and were repeated every 7 sec. The loudest tone in a series always had an intensity level of 110 db (re 10^{-16} watt/cm.^{2}), and the weakest an intensity level of 15 db. The tone was put through a narrow band-pass filter to eliminate switching transients. All of the sound-producing equipment was in a room adjacent to that of *O*. The tone production and timing were done with automatic electronic equipment. The intensities for each successive tone were produced by manual manipulation of an attenuator.

Experimental conditions.—A preliminary experiment was run on each *O* in which 20 stimulus categories spaced 5 db apart were used. The *O* was required to make a judgmental response to each tone heard on a scale from 1 to 20, with low numbers being used for weak tones and high

[1] This research was done in cooperation with the Systems Division, Naval Research Laboratory under Contract N5-ori-166, Task Order I, between the Office of Naval Research and The Johns Hopkins University. This is Report No. 166-I-171, Project Designation No. NR 507-470, under that contract. The author is indebted to Miss Frances Wolfram for assistance in the collection and analysis of the data.

numbers for loud tones; he was instructed to attempt to use each number for its corresponding tone. He was given an initial practice session in which each tone was identified for him with the appropriate number, and he was allowed to continue this practice until he felt that no further improvement would result.

From the data of this experiment an equal-discriminability scale (1, 2) of loudness was constructed for each *O* separately. This scale was used to construct the stimulus series for all other experiments, and all stimuli for the other experiments were spaced equidistant on the equal-discriminability scale. This procedure was used because maximum judgmental accuracy, or information transmission, should occur with such a spacing (2).

The main experiments varied only in respect to the total number of stimulus categories used. These were 4, 5, 6, 7, 10, and 20. For each experiment *O* was again given the practice session during which he tried to learn which numbers went with which loudnesses. The actual experimental series, then, were relatively free of practice effects.

Each experiment consisted of 100 presentations of each of the stimulus categories to each *O*. Thus the total number of judgments varied with the number of stimulus categories, ranging from 2,400 for 4 categories to 12,000 for 20 categories. For the experiments involving 10 and 20 categories, the series were set up so that successive pairs of stimuli occurred equally often, and no pair was repeated until every other pair had occurred at least once. Thus for the "20" experiment, a pair could not repeat until after 400 stimulus presentations, while a pair could repeat after 100 presentations for the "10" experiment. For the other experiments, the series were also set up so that every stimulus followed every other stimulus an equal number of times, but no restrictions about pairs occurring more frequently early or late in the series were used because of the small number of pairs.

Analysis of data.—The major analysis of data is in terms of information transmission, and for the purposes of this experiment it is convenient to think of the measure of information transmission as a measure of the predictability of the responses, given certain antecedent conditions. The methods of computing information transmission for the two-variable case (i.e., responses are predicted from just one variable) are given in Garner and Hake (2). The problem is more complicated when responses are predicted from more than one variable. A mathematical treatment of this type of analysis is available in McGill (6). In this paper, therefore, no attempt is made to explain completely all of the calculations which have been made.

Symbols.—Several terms referring to some of the variables analyzed are used frequently enough to warrant the use of symbols. The symbols used, and their meanings, are:

O—the observer.
S—the stimulus being judged on a particular presentation.
R—the judgmental response made by *O*.
P—the preceding stimulus. It is the stimulus which was presented just before the one which is being judged.
Q—the preceding response.

When these symbols are used, they refer to the appropriate variable, rather than to the specific event or object, except for *O*, which is used to refer both to the variable and to the person who is the observer.

Results

Number of categories.—The first analysis of judgmental accuracy was made by determining information transmission between *S* and *R* as a function of number of stimulus categories. The results of this calculation, for which data from all six *O*s were pooled, are shown in the lowest curve in Fig. 1. Essentially perfect accuracy was maintained with four or five categories. The information transmission in bits per stimulus presentation would be

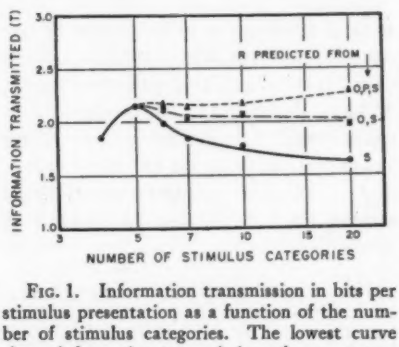


Fig. 1. Information transmission in bits per stimulus presentation as a function of the number of stimulus categories. The lowest curve shows information transmission when responses (*R*) are predicted only from stimuli (*S*), with data for all *O*s pooled. For the middle curve, *R* is predicted from *S* and from observer differences (*O*). For the top curve, *R* is predicted from *S*, *O*, and the preceding stimulus (*P*).

2.00 for four categories and 2.32 for five categories if no errors had been made. The obtained information transmissions are slightly less than these figures, and it is likely that this difference represents recording and scoring errors. For larger numbers of stimulus categories, however, judgmental accuracy becomes progressively poorer, being 1.62 bits for 20 categories. This figure is equivalent to perfect accuracy for only three categories. Thus the responses are much less predictable when the larger numbers of categories are used. Additional analyses were then made to determine why this loss of predictability occurred with the larger numbers of categories.

Observer effects.—Information transmission was next computed from a matrix in which *O* and *S* were considered as the input, with *R* as the output. The results of this calculation are shown in the middle curve of Fig. 1. It is clear from this analysis that a great deal of the decreased accuracy of judgment had to do with the fact that different *O*s were responding in different ways.

The information transmission measure used for this case is somewhat analogous to a multiple-correlation coefficient in which *R* is predicted from *O* and *S*. This total information transmission is equal to the information transmission between *S* and *R*, plus the information transmission between *O* and *R*, with *S* held constant—i.e., taking one *S* at a time, and then averaging over all *S*s. This latter term is equivalent to the partial correlation between *O* and *R*, with *S* constant. In other words, we can consider that there are two separately identifiable components of this information transmission, one resulting from *S* and the other from *O*. Thus the difference between the two lower curves in Fig. 1

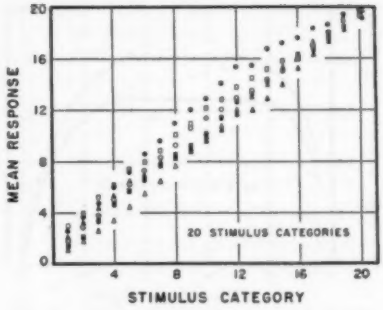


FIG. 2. Mean response as a function of stimulus category, in order of intensity. Each symbol is for a different *O*.

is a measure of the additional information transmission obtained from *O*s, or the extent to which knowledge of *O* increases the predictability of *R*.

The reason for this additional information transmission is shown in Fig. 2, where the mean response for each stimulus category is shown separately for the different *O*s. It is clear that all *O*s did not use the same response for a given stimulus category. For some there is simply a tendency to respond with higher or lower numbers on the average, while other *O*s show quite different shapes in their response functions. For example, the *O* graphed as a solid triangle shows responses which are near the highest used for low stimulus categories, and the lowest for high stimulus categories.

This fact of individual differences per se is less important than the fact that the additional information transmission obtained from knowledge of *O* is greater for the large numbers of stimulus categories than for the small numbers. No actual calculations were made for four and five stimulus categories because of the very few errors made, but it is clear that little or no improvement in transmission could have resulted.

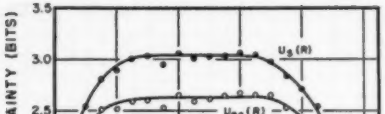


FIG. 3. Response uncertainty as a function of stimulus category, in order of intensity. The upper curve is response uncertainty for each stimulus category, regardless of preceding stimulus category (P). The lower curve represents the response uncertainty for one P at a time, and then averaged over all P's. The difference between these two curves is the additional information transmission obtained from knowledge of P.

Preceding stimulus.—Still further response predictability, or information transmission, can be obtained by considering combinations of S, O, and P (the preceding stimulus) as the input, with R still the output. Such a calculation would be equivalent to a multiple-correlation coefficient in which R is predicted from O, P, and S. In order to perform such a calculation, however, a matrix with 48,000 cells would be required when 20 stimulus categories were used, and the number of judgments available did not make such an analysis feasible. However, it is clear from the mathematical analysis that such a computation can be approximated by computing the information transmitted from P to R, with S held constant, and adding it to the information transmission from O and S to R. This term is equivalent to that which was added when we computed information transmission from O and S to R.

The addition of this term does not account for all of the information transmission, since there is an additional component involving combinations of O, S, and P. This unknown term can theoretically be negative, but it is clear that in this situation it would not be.

The top curve of Fig. 1 shows the combined information transmission from all three variables. It should be remembered that this curve is probably a slight underestimation of the true transmission values. It no longer shows any drop-off of information transmission with the larger number of categories. If anything, there is a slight rise. Once again the important point is that the additional information transmission gained from knowledge of the preceding stimulus is greater for the larger number of categories than for the smaller number.

The relation between predictability obtained from knowledge of P and S is shown in Fig. 3 and 4. Figure 3 shows response uncertainty as a function of stimulus category both when P is taken into account and when it is not for 20 stimulus categories. It is clear that the additional information about responses comes primarily from stimuli in the middle categories, where response uncertainty is lowered most. Figure 4 shows the same two curves when 6 stimulus categories were used. Although the curves are similar to

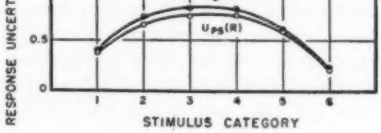


FIG. 4. Response uncertainty as a function of stimulus category, in order of intensity. The two curves have the same meanings as those in Fig. 3.

those in Fig. 3, the additional information transmission obtained from knowledge of P is very slight.

The over-all shapes of the curves shown in Fig. 3 and 4 are of some interest. Response uncertainty is least at the two ends of the stimulus continuum. This is because errors can be made in only one direction at the ends, so that the distribution of responses is much narrower. In addition, there seems to be less uncertainty in the responses to the higher stimuli than to the lower stimuli. This effect is an artifact of the tendency of *O*s to use the higher numbers somewhat more often than they should, as can be seen in Fig. 2, since the more often *O* uses a particular number, the more apt he is to be correct when the stimulus category equivalent to that number is presented. The same tendency, however, results in more uncertainty about which stimulus was presented when we know the response. Figure 5 shows this effect when 20 stimulus categories were used. Here it is clear that the lower numbers are used with somewhat greater accuracy, although they are used less often than the higher numbers.

Figures 6 and 7 show the nature of the effect which produces the additional information transmission when P is taken into account. From Fig. 6, for 20 stimulus categories, it can be seen that there is a tendency for a response to a given stimulus category to be higher when the preceding stimulus is higher than the one being judged, and lower when the preceding stimulus is lower. Although the effect is diminished with only 6 stimulus categories, as shown in Fig. 7, it still exists.

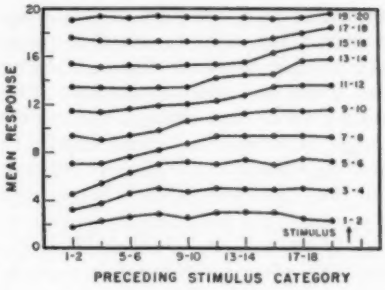


FIG. 6. Mean response as a function of preceding stimulus category, for various stimulus categories. Pairs of stimulus and preceding stimulus categories have been averaged for these functions. Data from all *O*s pooled.

Distribution of stimuli.—It was mentioned in a previous paper (2) that maximum information transmission should occur if the stimuli are spaced

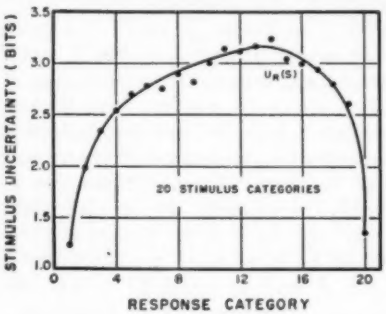


FIG. 5. Stimulus uncertainty as a function of response category.

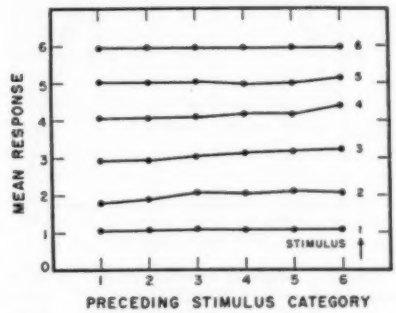


FIG. 7. Mean response as a function of preceding stimulus category, for various stimulus categories. Data from all *O*s pooled.

according to a criterion of equal discriminability, and in another paper (1) a scale of equal discriminability for loudness was developed. This scale showed that through the middle range of intensities, discriminability and log intensity are linearly related, but that discriminability is better near the ends of the judgmental continuum than in the middle. Pollack (8) found the same effect for judgments of pitch, although he did not develop an actual discriminability function.

The data of Fig. 1 were all obtained with stimuli spaced along a scale of equal discriminability. This scale was obtained for each *O* with the preliminary experiment in which stimuli were spaced in 20 equal db steps. A comparison of the results of this experiment with that using 20 equal discriminable steps allows us to check the assumption of better information transmission with the latter. Figure 8 shows response uncertainty as a function of stimulus category for this preliminary experiment, both when *P* is taken into account and when it is not. Comparison of these curves with those of Fig. 3 shows the effect of the different spacing of the stimuli. For the curves of Fig. 3, more stimuli were crowded into the ends of the continuum and fewer in the middle; as we would expect, response uncertainty became greater at the ends and better in the middle compared to the curves of Fig. 8. The gain in predictability of response to stimuli in the middle categories has more than offset the loss at the ends, so that over-all predictability is better.

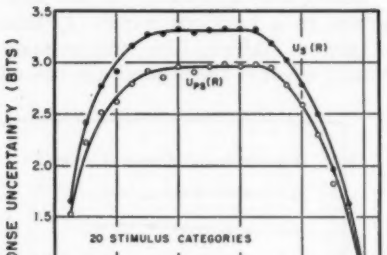


FIG. 8. Response uncertainty as a function of stimulus category, in order of intensity. These curves and their meanings are the same as those of Fig. 3, except that the stimuli have been spaced in equal db steps, rather than equal discriminable steps.

When data for all *O*s are pooled, the information transmission between *S* and *R* is 1.53 bits for equal db steps, and 1.62 for equal discriminable steps. When account is also taken of both *P* and *O*, the comparable figures are 2.16 and 2.29 bits. The difference between these figures is not large, but it seems likely that better transmission does occur when equal discriminable steps are used. It is, of course, not possible to demonstrate that this condition is optimal.

Other factors.—There are still other factors which affect information transmission, or the predictability of the responses, which cannot be handled precisely in these experiments because of the impossibility of doing a complete analysis with so many categories. For example, the preceding response (*Q*) affects responses in much the same way as *P*. If curves of Fig. 6 and 7 are drawn, with *Q* used rather than *P*, they are very similar, but slightly steeper. It is obvious that if there is some predictability from *P*, there must also be from *Q*, since these are so highly correlated. It is also clear from inspection of the data that there would be even more predictability if both *P* and *Q* could be taken into account together, but an exact analysis is impossible.

Another factor which influences pre-

dictability of responses is the experimental session. For example, when 20 stimulus categories were used, data had to be collected on at least three different days. The *O*'s criteria for the different response categories can change from day to day, and even from session to session of the same day. In order to estimate the size of this effect, the data for 20 categories were broken into five groups, and those for 10 categories into two groups. Then information transmission between groups and *R* with *O* and *S* held constant was computed. This figure was .22 bit for 20 categories and .05 bit for 10 categories. These figures cannot be added onto the already computed information transmission with complete assurance because of possible interactions with some of the other variables. It is interesting to note, however, that this additional predictability of responses is considerably greater for the larger numbers of stimulus categories than for the smaller numbers.

Discussion

Effect of preceding stimulus.—Figures 6 and 7 showed that the response to a stimulus is higher when the preceding stimulus was higher than that judged, and lower when the preceding stimulus was lower than that judged. This effect has been noted before (5) in a situation in which a standard stimulus occurred before each judged stimulus. This result is contrary to what would be expected from Michels and Helson's (7) formulation of the concept of judgments with respect to an adaptation level, since this formulation would require that a high preceding stimulus raise the adaptation level, and thus lower the response to this particular stimulus with respect to that adaptation level.

This effect is also related to many of the time error phenomena, which are thoroughly discussed in Woodworth (9). For example, the effect noted here will produce an over-all "central tendency" of judgment because of the statistical fact that in a given series more high than low tones will precede low tones, and conversely for high tones. Thus on the average there will be a tendency to respond with relatively too many middle values, or to show a regression of responses toward the mean response. However, it is clear from the analysis shown here that this tendency is not as strong with respect to the middle of the series as it is with respect to what has just been heard. In a sense, what has just been heard becomes the "center" toward which responses are made.

Actually, there are two possible explanations of the effect of the preceding stimulus on the judgment, neither of which can be conclusively demonstrated without further analysis than is possible here. The first is that the response to a stimulus is actually a response to the weighted mean of the present stimulus and that heard previously. Thus the response will always be biased high or low, depending on what has recently been heard. The other possible explanation is that there is a tendency on the part of *O* to repeat what he has just said when he is unsure of his judgment. The difference between these two possibilities is that in the first case what has been heard influences the present response, while in the latter what *O* has said influences the present response. In order to separate these two possible effects, it will be necessary to get an analysis of the mean response as a function of combinations of *S*, *P*, and *Q* jointly.

Judgmental complexity.—We can think of the differences in number of stimulus categories as differences in the complexity of the judgmental situation. As more stimulus categories are added, the discriminability between adjacent stimuli becomes poorer, there are more response categories to use, and it becomes much more difficult for *O* to keep just one response number associated with one stimulus category.

The results of these experiments have shown that, if we know only the stimulus, we can more accurately predict the

response in the simpler judgmental situation than in the more complex one. However, it has also been shown that the predictability of responses is not essentially different between the simpler and more complex situations if a sufficient number of predictive factors is taken into account. Thus the important point is that in the complex judgmental situation, the response is dependent on a great many factors other than the stimulus presented for judgment, while in the simpler situation, these other factors have very little effect on the judgment.

A summary of the effect of these other factors when 20 stimulus categories were used will illustrate how great this difference is. The information transmission between *S* and *R* is 1.62 bits, which is equivalent to just over three categories. When the increased predictability from *P* and *O* is taken into account, the information transmission has risen to 2.29 bits, which is equivalent to approximately five categories. If we add the predictability that results from taking account of differences between sessions, this figure increases to 2.51 bits, which is approximately six categories. Thus taking account of factors other than *S* has essentially doubled the number of categories which could be used with perfect accuracy. Although there is some error in the estimation of these figures, the error is probably in the direction of underestimating the number of categories.

Summary

These experiments required *O*s to make absolute judgments of the loudness of tones, using numbers corresponding to the number of stimulus categories, which varied from 4 to 20 in the different experiments. The primary analysis was of information transmission between various factors and the responses. The major results and conclusions are:

1. If only stimuli are used to predict responses, maximum information transmission occurs with five stimulus categories and decreases with more stimulus categories. Thus responses will be most accurately related to stimuli when relatively few categories are used.

2. The predictability of responses with the larger number of stimulus categories is as great as, or greater than, it is with the smaller number if factors other than the stimulus are used to predict the response. Some of these factors are differences between *O*s, previous responses and stimuli, experimental session, and distribution of stimuli on the intensity continuum.

3. The effect of the preceding stimulus on responses is to raise the response if the preceding stimulus is higher than the one being judged, and to lower it if it is lower than the one being judged. Since a similar effect of preceding responses exists, it is not clear whether the effect is due to *O*'s responding to a weighted average of the present and preceding stimulus or to a tendency to repeat the same response.

References

1. Garner, W. R. An equal discriminability scale for loudness judgments. *J. exp. Psychol.*, 1952, **43**, 232–238.
2. Garner, W. R., & Hake, H. W. The amount of information in absolute judgments. *Psychol. Rev.*, 1951, **58**, 446–459.
3. Hake, H. W., & Garner, W. R. The effect of presenting various numbers of discrete steps on scale reading accuracy. *J. exp. Psychol.*, 1951, **42**, 358–366.
4. Klemmer, E. T., & Frick, F. C. Assimilation of information from dot and matrix patterns. *J. exp. Psychol.*, 1953, **45**, 15–19.
5. Long, L. A study of the effect of preceding stimuli upon the judgment of auditory intensities. *Arch. Psychol., N. Y.*, 1937, **30**, No. 209.
6. McGill, W. J. Multivariate transmitted information. *Psychometrika*, in press.
7. Michels, W. C., & Helson, H. A reformulation of the Fechner law in terms of adaptation-level applied to rating scale data. *Amer. J. Psychol.*, 1949, **62**, 355–368.
8. Pollack, I. The information of elementary auditory displays. *J. acoust. Soc. Amer.*, 1952, **24**, 745–749.
9. Woodworth, R. S. *Experimental psychology*. New York: Holt, 1938.

(Received for early publication July 13, 1953)

PSYCHOLOGICAL MONOGRAPH

#324
1951

A recent psychological monograph of general interest

VISUAL PERCEPTION AND THE ROTATING TRAPEZOIDAL WINDOW

BY

ADELBERT AMES, JR.

$1.00

Order from

American Psychological Association
1333 Sixteenth Street N.W.
Washington 6, D. C.

PSYCHOLOGICAL MONOGRAPH

#270

1945

ON PROBLEM SOLVING

By

KARL DUNCKER

$2.50

This popular monograph is #270 of the Psychological Monograph series. It has been reprinted so that it is again available.

Third Printing

AMERICAN PSYCHOLOGICAL ASSOCIATION

1333 Sixteenth Street N. W.

Washington 6, D. C.